

JOURNAL Submitted (16-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Jul 12, 2001 this sequence version replaced gi:12654174.
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Huliy, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 3 Row: k Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4886442.

FEATURES
source Location/Qualifiers

1..1433
/organism="Homo sapiens"
/db_xref="locusid:5861"
/db_xref="taxon:9606"
/clone="MGC:5233 IMAGE:2900705"
/tissue="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
179..796
/codon_start=1
/product="RAB1, member RAS oncogene family"
/protein_id="AAH0905.1"
/db_xref="gi:12654175"
/translation="MSEMPEDYDLFKLLIGDSGVKSCLLRFADDTYESTISIT
KQDFKRTIELDKRTIKLOIWDPAQOERPTTSSVYRGHGLIIVYDVPDPSFNV
KQWLEIDRASENVKLLVGNKCDLTKKVVDTTAKERADSLGIDFLETSKMNATN
VEQSEFMMAEIKRMKPGATAGAAKSNVKIQTSPVKQSGGSCC"

CDS
BASE COUNT 432 a 254 c 326 g 421 t

ORIGIN
Query Match 86.8% Score 1219: DB 9: Length 1433;
Best Local Similarity 93.2% Pred. No. 2.8e-187;
Matches 1336; Conservative 0; Mismatches 0; Indels 97; Gaps 2;

QY 3 GCGATAGCTGAGT-GCGGCGGCTGCTGATTTGTTCTAGGAGGAGTAGGGAAGACG 61
Db 1 GCGATAGCTGAGTGGCGGCGGCTGCTGATTTGTTCTAGGAGGAGTAGGGAAGACG 60
QY 62 TTTGCTCTCCCGGAACACCTATCTCATCTCTTCTTCGATTACCCGTGCGGAGAG 121
Db 61 TTTCTCTCTCCCGGAACACCTATCTCATCTCTTCTTCGATTACCCGTGCGGAGAG 120
QY 122 TCAGGGGGGGGGGCGGACAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 181
Db 121 TCAGGGGGGGGGGCGGACAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 180
QY 182 GTCCAGCATGATCCGATATGATTTATTTATCAAGTTACTTCTGATTGGCAGTCAAG 241
Db 181 GTCCAGCATGATCCGATATGATTTATTTATCAAGTTACTTCTGATTGGCAGTCAAG 240
QY 242 GGTGGAAGTCTTGCTTCTTTAGGTTTGAAGATGATATACATATACAGAAAGTCAAT 301
Db 241 GGTGGAAGTCTTGCTTCTTTAGGTTTGAAGATGATATACATATACAGAAAGTCAAT 300
QY 302 CAGCACAATTGGTGTGATTTCAAAATTAAGACTATAGAGTTAAGCGGAAACATCA 361
Db 301 CAGCACAATTGGTGTGATTTCAAAATTAAGACTATAGAGTTAAGCGGAAACATCA 360

Db 301 CAGCACAATTGGTGTGATTTCAAAATTAAGACTATAGAGTTAAGCGGAAACATCA 360
QY 362 GCTTCAATA----- 371
Db 361 GCTTCAATATAGGACACAGCAGCGCAGGAAGATTTCGAAATCACTCCACTTATTA 420
QY 372 -----GAGTCTTCAATTA 385
Db 421 CAGAGAGCCCATGACATCATAGTTGTATGATGTGACAGATCAGAGTCCCTCAATTA 480
QY 386 TGTTAACAGTGGTGCAGAGAAATAGATCCGTATCCAGTCAAAATGTCAAATGTT 445
Db 481 TGTTAACAGTGGTGCAGAGAAATAGATCCGTATGCGAGTCAAAATGTCAAATGTT 540
QY 446 GGTAGGGAACAATGTATCTGACACAAAGAGTAGATACATCAACAGCGAAGGA 505
Db 541 GGTAGGGAACAATGTATCTGACACAAAGAGTAGATACATCAACAGCGAAGGA 600
QY 506 ATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTCTAAGATGCAACGAATGT 565
Db 601 ATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTCTAAGATGCAACGAATGT 660
QY 566 AGAACAGTCTTCAATGACAGATGCGAGCTGAGATTAAAGAGGATGGTCCCGAGGAC 625
Db 661 AGAACAGTCTTCAATGACAGATGCGAGCTGAGATTAAAGAGGATGGTCCCGAGGAC 720
QY 626 AGCTGTGCTGCTGAGAAGTCCAAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 685
Db 721 AGCTGTGCTGCTGAGAAGTCCAAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 780
QY 686 TGAAGTGTCTGCTAAATTTGCTCCATCTTTTTCACAGCAATGAATTTGCAATCTG 745
Db 781 TGAAGTGTCTGCTAAATTTGCTCCATCTTTTTCACAGCAATGAATTTGCAATCTG 840
QY 746 AACCAGTGAAGAAACAAATTCCTGAATGCTGATGATGATGCTGCACTACACAGA 805
Db 841 AACCAGTGAAGAAACAAATTCCTGAATGCTGATGATGATGCTGCACTACACAGA 900
QY 806 TTCTTACGCTCCACAAAGTGCAGAGATGTAATGCTCAATGCTACTGCTTTTAT 865
Db 901 TTCTTACGCTCCACAAAGTGCAGAGATGTAATGCTCAATGCTACTGCTTTTAT 960
QY 866 TCCCTTGACACAGACGCTACTTCAATTTTCAAGACTGTTTAAACCTTTGTGCTGG 925
Db 961 TCCCTTGACACAGACGCTACTTCAATTTTCAAGACTGTTTAAACCTTTGTGCTGG 1020
QY 926 TTTTAAATATATGATGATTCCTTGTGCTTCCGATGACGAGCTGTTTCCGCTGT 985
Db 1021 TTTTAAATATATGATGATTCCTTGTGCTTCCGATGACGAGCTGTTTCCGCTGT 1080
QY 986 TGGTTAGATATATTTGTTTGTATGTTTATATTTGATGATGTCAGGTTTATGTC 1045
Db 1081 TGGTTAGATATATTTGTTTGTATGTTTATATTTGATGATGTTTATGATGTCAGGTTTATGTC 1140
QY 1046 TTTGGAAGATTAATTCAGCATTTTGTATTAACACAGCAGAGAGTGTGCTACCTTC 1105
Db 1141 TTTGGAAGATTAATTCAGCATTTTGTATTAACACAGCAGAGAGTGTGCTACCTTC 1200
QY 1106 CATGATTAATTTAGTGAATGTTATATGTAATGATGATGATTTGCTAGTCTTCTTCTGTA 1165
Db 1201 CATGATTAATTTAGTGAATGTTATATGTAATGATGATGATTTGCTAGTCTTCTTCTGTA 1260
QY 1166 GAGTTATTAATGAAAGATTAACATATCTGATTAATGATTTCTTCAATCTGATTA 1225
Db 1261 GAGTTATTAATGAAAGATTAACATATCTGATTAATGATTTCTTCAATCTGATTA 1320
QY 1226 TTTGTGCTGACAGATTTGTAATTTGTGACACCTATGTAACAAACCACTGAAGTAT 1285
Db 1321 TTTGTGCTGACAGATTTGTAATTTGTGACACCTATGTAACAAACCACTGAAGTAT 1380
QY 1286 GTTAAATTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 1338
Db 1381 GTTAAATTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 1433

RESULT 2
AK055927 2257 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens CDNA FLJ131365 fis, clone NB9N41000135, highly similar
DEFINITION to RAS-RELATED PROTEIN RAB-1A.
ACCESSION AK055927
VERSION AK055927.1 GI:16550777
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens neuroblastoma cell line:NB9 CDNA to mRNA,
clone lib:NB9N41 clone:NB9N41000135.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 Tashiro, H., Yamazaki, M., Matanabe, K., Kumagai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Nakamatsu, A., Ishi, S., Yamamoto, J., Isono, Y.,
Kawai-Hiro, Y., Saito, K., Nishikawa, T., Kikuchi, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, K., Kanda, K.,
Wagatsuna, M., Murakawa, K., Kanedoki, K., Takahashi, Fujii, A.,
Oshiba, A., Sugiyama, A., Kawakami, B., Suzuki, T., Sugano, S.,
Nagahara, K., Masuno, T., Nagai, K., and Sugiyama, T.
TITLE NBO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2257)
AUTHORS Isogai, T., Otsuki, T., and Sugiyama, T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: CDNA full insert sequencing:
Research Association for Biotechnology (RAB): CDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'- end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
FEATURES
source Location/Qualifiers
1..2257
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NB9N41000135"
/cell_line="NB9"
/cell_type="neuroblastoma"
/clone_lib="NB9N41"
/note="cloning vector: pME18FL3"
194..619
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71048.1"
/db_xref="GI:16550778"
/translation="MSNNPEYDYLFLKLLIGDSGVGSKLLLEAFESNNYKWLQIE
IDRYASNNVKLLVGNKCDLTKKVVVDYTKAFEDSLGIFLETSAKNATNVEDSPM
TMAEIKRMGPGATFAGAGSEKSNYKTIQSTPVKSGGCC"
BASE COUNT 669 a 375 c 482 g 731 t
ORIGIN
Query Match 78.5% Score 1103.6; DB 9; Length 2257;
Best Local Similarity 92.4%; Pred No.1-2e-168;
Matches 1223; Conservative 0; Mismatches 4; Indels 97; Gaps 2;
QY 1 AAGGATAGCTGAGT-GGGCGCGCTGCTGATGCTGTTTACGAGGAGCGATGGGGAAGA 59
DB 14 AAGGATAGCTGAGTGGCGCGCTGCTGATGCTGTTTACGAGGAGCGAGGGAAGA 73
QY 60 CGTTGCTGCTCCGGAAGACCTTATCTCATCTCTTCTTGGATACCCGTTGGCGGAG 119
DB 74 CGTTGCTGCTCCGGAAGACCTTATCTCATCTCTTCTTGGATACCCGTTGGCGGAG 133
QY 120 AGTAGAGGCGCGCGCTGCGGAGGAGGCGCGCTGCGGCGGCGGAGCTGACAGTGAC 179

DB 134 AATCAGGCGCGCGCTGCGGAGGAGGCGCGCTGCGGCGGCGGAGCTGACAGTGAC 193
QY 180 ATGTCAGCATGAATCCCGAATATGATTTATTTATTCAGTTACTTGTGTTGGGACTCA 239
DB 194 ATGTCAGCATGAATCCCGAATATGATTTATTTATTCAGTTACTTGTGTTGGGACTCA 253
QY 240 GGGGTGGGAATGCTTCGCTTCCTTGTAGTTTGCAGATGATTCATATTCAGGAAGCTAC 299
DB 254 GGGGTGGGAATGCTTCGCTTCCTTGTAGTTTGC----- 288
QY 300 ATCAGCATGATGTTGTTGATTTCAAAATTAAGAACTATAGATTAGAGGGGAAAAATC 359
DB 289 ----- 288
QY 360 AAGCTCAATATGATGCTTCATATTTATTAACAGTGGCTGAGGAATATGATCGTTAT 419
DB 289 ----- 337
QY 420 GCCAGTGAATATGTCACAAATTTGTTGAGGGAACAAATGTGATCTGACCAAGAA 479
DB 338 GCCAGTGAATATGTCACAAATTTGTTGAGGGAACAAATGTGATCTGACCAAGAA 397
QY 480 GTAGTAGCTACACAACGGAAGGAATTTGCTATTTCCCTTGGAATTCGTTTGGAA 539
DB 398 GTAGTAGCTACACAACGGAAGGAATTTGCTATTTCCCTTGGAATTCGTTTGGAA 457
QY 540 ACCAGTGTAAAGATGCAACGAATGTAGAACAGTCTTTCATGACGATGGCAGTGAGAT 599
DB 458 ACCAGTGTAAAGATGCAACGAATGTAGAACAGTCTTTCATGACGATGGCAGTGAGAT 517
QY 600 AAAAAGCGAATGGGTCCCGGAGGAGCAACAGTGGTGGTCTGAGAGTCCAAATGTTAAAT 659
DB 518 AAAAAGCGAATGGGTCCCGGAGGAGCAACAGTGGTGGTCTGAGAGTCCAAATGTTAAAT 577
QY 660 CAGAGCATCCAGTCAGCAGTCAAGTGGAGTTGCTGCTAAATTTCCCTCATCTCTT 719
DB 578 CAGAGCATCCAGTCAGCAGTCAAGTGGAGTTGCTGCTAAATTTCCCTCATCTCTT 637
QY 720 TCTCAGCAATGAATTTGCAATCTGACCAAGTGAAGAAACAAATTTGCTCAATTTCT 779
DB 638 TCTCAGCAATGAATTTGCAATCTGACCAAGTGAAGAAACAAATTTGCTCAATTTCT 697
QY 780 ACTGTATGTAGTGAATGACAGATGCTTACCGTCCGACCAAGGTCAGAGTTGTAA 839
DB 698 ACTGTATGTAGTGAATGACAGATGCTTACCGTCCGACCAAGGTCAGAGTTGTAA 757
QY 840 ATGTCATATGACGATTTTATTTATTCCTTGAACAGAGTCACTACTCTTTTCAG 899
DB 758 ATGTCATATGACGATTTTATTTATTCCTTGAACAGAGTCACTACTCTTTTCAG 817
QY 900 AACTGTTTAAACCTTTGCTGCTGTTTAAATTAATGCTGAATCTGCTGCTT 959
DB 818 AACTGTTTAAACCTTTGCTGCTGTTTAAATTAATGCTGAATCTGCTGCTT 877
QY 960 CCGATACAGACGCTTCCCGGCTGCTGTTAGAAATTAATTTGTTGATGTTTATAT 1019
DB 878 CCGATACAGACGCTTCCCGGCTGCTGTTAGAAATTAATTTGTTGATGTTTATAT 937
QY 1020 GCGATGTTAAATGTCAGGTTTGTCTTCTGAAGTGAAGTCACTTGTGATCAAA 1079
DB 938 GCGATGTTAAATGTCAGGTTTGTCTTCTGAAGTGAAGTCACTTGTGATCAAA 997
QY 1080 CAGCAGACAGAGTGTCTGCTCACTTCCATGATTAAGTTAGTGAATGATTAATGTAAG 1139
DB 998 CAGCAGACAGAGTGTCTGCTCACTTCCATGATTAAGTTAGTGAATGATTAATGTAAG 1057
QY 1140 ATCTGATTTGCTAGTTCTTCTGTTAGAGTTATTAATGGAAGATTTACATCTGATTA 1199
DB 1058 ATCTGATTTGCTAGTTCTTCTGTTAGAGTTATTAATGGAAGATTTACATCTGATTA 1117
QY 1200 ATAGTTTCTCATCTCTGATATTAATTTGTGCTGAGAAATTTGTAATTTTTCACA 1259

Db	1118	ATAGTTTCTCTACTCTGCANATATATTGTGGCGCGAAGANATGTGAATTGTTCGACA	1177
QY	1260	CTATGTAAACAAACACGTGACAGATATGTTTAATAATATTTGCTATTGTGGAAGTA	1319
Db	1178	CTATGTAAACAAACACGTGAGATATGTTTAATAATATTTGACTATTGGAAGTAATAT	1237
QY	1320	AAAA	1323
Db	1238	CAAA	1241
RESULT 3			
BC002077		1444 bp	RNA linear
LOCUS	BC002077		
DEFINITION	Mus musculus, RAB1, member RAS oncogene family, clone MGC:6226		
ACCESSION	IMAGE:3592802, mRNA, complete cds.		
VERSION	BC002077		
KEYWORDS	BC002077.1	GI:12805232	
SOURCE	MGC.		
ORGANISM	house mouse.		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 1444)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JAN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.shgc.stanford.edu Contact: (Dickson, Mark) mcdspaxil@stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 11 Row: a Column: 2 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6679586.		
source	Location/Qualifiers		
	1. 1444		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/map="FVB/N"		
	/clone="MGC:6226 IMAGE:3592802"		
	/tissue_type="Mammary tumor. Metallochonien-TGF alpha model. 10 month old Virgin mouse. Taken by biopsy."		
	/clone_id="NCI CGAP_Mam1"		
	/lab_host="DH10B"		
	/note="Vector: pCMV-SPORT6"		
	196. 813		
	/codon_start=1		
	/product="RAB1, member RAS oncogene family"		
	/protein_id="AAH02077.1"		
	/db_xref="GI:12805233"		
	/db_xref="locusid:5861"		
	/translation="MSSKNPFDYDLFKLLIGDSGVKSCLLFRADDYTESITITIGVDKRIELDTGKTIKLQMDTAGDGEFRITSSYKGAAGIIVYDVDDSEFNANVGKQMDIEDRYASENVNKLQVGNKCDLTKTKVDDYTAKEFADSDIGIFLETSANANNAVEOSFMTAAEIKRKMPGATAGGAEKSNVNIQSTPVKSGSGCC"		
CDS			
BASE COUNT	412 a	270 c	342 g 420 t
Query Match	74.9%	Score 1052.6;	DB 10; Length 1444;

Best Local Similarity 88.3%, Pred. No. 1,9e-160:	
Matches 1256; Conservative 0; Mismatches 59; Indels 108; Gaps 6:	
QY	16 GCGGCGGCTGCTGATGTGTCTTAGGGAGCGAGTGGGGAAAGAGTTTGCTCTCCGGA 75
Db	27 GCGGCGGCTGCTGATGTGTCTTAGGGAGCGAGTGGGGAAAGAGTTTGCTCTCCGGA 85
QY	76 ACAGCTTATCTATCTCTTTCTTTGATTACCCGTGGCGCGGAGTACAGGGCGCGCT 135
Db	86 ACAGCTTATCTATCTCTCTTTCTTTGATTACCCGTGGCGCGGAGTACAGGGCGCGCT 145
QY	136 GCGGAGCAAGGCG-----GGCGGTGGCGGGCGGGAGCTGCAGTACATGTGCACA 189
Db	146 GCGTACGCAAGGCGCGGTGTGGCGGGCGGGAGCTGCAGTACATGTGCACA 205
QY	190 TGAATCCCGAATGATGATTTATTTTCAAGTTACTCTGATTGGCGACTCAGGGGTTGAA 249
Db	206 TGAATCCCGAATGATGATTTATTTTCAAGTTACTCTGATTGGCGACTCTGGGGTTGAA 265
QY	250 AGCTTGCCCTTCTTTAGGTTTGCAAGTATATACATATACAGAAAGCTATCAGACAA 309
Db	266 AGCTTGCCCTTCTCTTAGGTTTGCAAGTATATACATATACAGAAAGCTATCAGACAA 325
QY	310 TTGGTGGGATTTCAAAATAGAACTATAGATTAGAGGGGAAACATTCAGCTTCAA 369
Db	326 TTGGTGGGATTTCAAGATACGAACTATAGATTAGATGGGAAACATTCAGCTTCAA 385
QY	370 TA----- 371
Db	386 TATGGGACACAGCAGGCGGAGAAATTTGCAACAATCACTTCAGTTATTACAGAGAG 445
QY	372 -----GAGTCTTCAATTAATGTTAAAC 393
Db	446 CCCATGGCATCATGATTGTGTATGATGTGACAGATCAGGAGTCTTTCATATAACGTTAAAC 505
QY	394 AGTGCTCGAGGAATAGATGCTTATGCGCAGGAAATGTCAACAAATGTTGGTAGGA 453
Db	506 AGTGCTCGAGGAGATAGTATGCTACGCGCAGTGAATATGTCAACAAGTTGTTGGTAGGA 565
QY	454 ACAAAATGTATCTGACCAACAAGAAAGTAGTAGACTACACAACAGCAGGAAGTAATTTGCTG 513
Db	566 ACAAAATGTAGACTGACCAACAAGAAAGTAGTAGACTACACAACAGCAGGAAGTAATTTGCGAG 625
QY	514 ATTCCTTGGAATTCCTTTTGGAAACAGTGTCTAAGATGCAACGAATGTAGAACAGT 573
Db	626 ATTCCTTGGAATTCCTTTTGGAAACAGTGTCTAAGATGCAACGAATGTAGAACAGT 685
QY	574 CTTTATGACGATGGACGCTGATTAATAAACGAATGGGTCCGAGACAAAGCTGGTG 633
Db	686 CTTTATGACGATGGACGCTGATTAATAAACGAATGGGTCTTGAGCTTACAGCTGGTG 745
QY	634 GTGCTGAGAAAGCCATGTTAAATTCAGAGCACTCCAGTCAAGCAGTCAAGTGGAGTT 693
Db	746 GTGCGAGAGATGCCATGTTAAATTCAGAGCACTCCAGTCAAGCAGTCAAGTGGAGTT 805
QY	694 GCTGCTAAATTTGCTCTCATCTTTTCTCACAGCAATGATTTGCAATCTGACCCAG 753
Db	806 GCTGCTAAATTTGCTCTCATCTTTTCTCACAGCAATGATTTGCAATCTGACCCAG 865
QY	754 TCAAAAAACAAATTGCTCGAATTTTACTGTATGATGAGTGCATCAACAGATTTCTTACC 813
Db	866 TCAAAAAACAAATTTGCTCGAATTTTACTGTATGATGAGTGCATCAACAGATTTCTTACC 925
QY	814 GTGCTCACAAAGGTAGAGATGTAAATGGTCAATACTGAC--TTTTTTTTTATTTCCCTTG 872
Db	926 GTTTCCAC--AAGGTAGAGATGTAAATGGTCAATACTGACATTTTTTTTTTATTTCCCTTG 984
QY	873 ACTCAAGACAGTAACTTCTATTTTCAGAACTGTTTAAACCTTTGTGCTGGTTATAA 932
Db	985 ACTCAAGACAGTAACTTCTATTTTCAGAACTGTTTAAACCTTTGTGCTGGTTATAA 1044
QY	933 AATTAATGTGTAACTCTTGTGCTTCTCTGATACAGACTGTTTCCGCTGGTTGATTAG 992

Db 1045 AATAATGCTGAATCCCTGCTTCTTCCTTCGATACCAAGATCCCTTCCCGGTGGTGTAG 1104
QY 993 AATAATTTTGTGTTGATGTTTATATATGCGATGTTTATGATGTTTATGCTCTGTGA 1052
Db 1105 AATAATTTTGTGTTGATGTTTATATGCGATGTTTATGATGTTTATGCTCTGTGA 1164
QY 1053 GATGAAGTTCAGCCATTTTGTATCAACAGACAGACAGTGTCTGTCACTTTCATCAT 1112
Db 1165 GATGAAGTTCAGCCATTTTGTATCAACAGACAGACAGTGTCTGTCACTTTCATCAT 1224
QY 1113 AAGTTTGTAGATGTTTATGATGAATGTTTATGATGTTTATGATGTTTATGATGTTTAT 1172
Db 1225 AAGTTTGTAGATGTTTATGATGAATGTTTATGATGTTTATGATGTTTATGATGTTTAT 1284
QY 1173 AATGGAAGATTCACATATCTGATTAATGTTTATGATGTTTATGATGTTTATGATGTTTAT 1232
Db 1285 AATGGAAGATTCACATATCTGATTAATGTTTATGATGTTTATGATGTTTATGATGTTTAT 1344
QY 1233 CTGCAATATTTGATTTGTTGACACTATGTATGACAAACAGTGAATGATGTTTAT 1292
Db 1345 CTGCAATATTTGATTTGTTGACACTATGTATGACAAACAGTGAATGATGTTTAT 1401
QY 1293 AATATTTGCTTATTTGGAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1335
Db 1402 AATATTTGCTTATTTGGAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1444

RESULT 4
LOCUS HSM800023 1208 bp mRNA linear PRI 10-MAR-2001
DEFINITION Homo sapiens mRNA; cDNA DKFZp564B163 (from clone DKFZp564B163);
ACCESSION AL050268
VERSION AL050268.1 GI:4886442
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1208)
AUTHORS Wiemann,S., Well,B., Wellenreuther,R., Gassenhuber,J., Glassl,S.,
Ansoerger,W., Boecher,M., Bloecher,H., Bauersachs,S., Blum,H.,
Lauder,J., Duesterhoeft,A., Beyer,A., Koehler,K., Strack,N.,
Mewes,H.W., Ottenwaelder,B., Obermaier,B., Tempe,J., Heubner,D.,
Wambutt,R., Korn,B., Klein,M. and Poustka,A.
Toward a Catalog of Human Genes and Proteins: Sequencing and
Analysis of 500 Novel Complete Protein Coding Human cDNAs
Genome Res. 11 (3), 422-435 (2001)
JOURNAL 11230166
PUBMED 2 (bases 1 to 1208)
REFERENCE Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
AUTHORS Direct Submission
TITLE Submitted (10-MAR-1999) MIPS, Am Klopferstr 18a, D-82152
JOURNAL Martinstr.10, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZ564B163) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY. Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
FEATURES
source location/Qualifiers
1..1208
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp564B163"
/issue_type="brain"
/clone_lib="564 (synonym: hfb2). Vector pAMP1; host
XL-2Blue; sites NotI + SalI"
/dev_stage="fetal"
191..580
gene

CDS
/gene="DKFZp564B163"
191..580
/gene="DKFZp564B163"
/note="Rab1, splice variant"
/codon_start=-1
/product="hypothetical protein"
/protein_id="CAB43369.1"
/db_xref="GI:4886443"
/db_xref="SPTREMBL:O9372"
/translation="MSSNPEDYLFILIGDSGVKSCLLRFADDTYESYTI
GVDFIRITELDKTKLIDFADSLGIPLETISAKNATNENQSFVMAEIKRMGP
GATGACAEKSNVRIQSTPVKOSGGGCC"
1169..1174
polya_signal
polya_site 1189
BASE COUNT 344 a 213 c 279 g 372 t
ORIGIN
Query Match 74.3% Score 1044; DB 9; Length 1208;
Best Local Similarity 90.0%; Pred. No. 4,7e-159;
Matches 1197; Conservative 0; Mismatches 0; Indels 133; Gaps 2;
1 AAGGATAGCTGACT- GCGCGCGCTGCTGATTTGTTTACGGGACGGAGTACGGAAGA 59
Db 11 AAGGATAGCTGAGTGGCGGCGCTGATTTGTTTACGGGACGGAGTACGGAAGA 70
QY 60 GCTTGTCTCCCGGAGACGCTATCTCATCTCTTCTTTCGATTAACCGTGGCGGAG 119
Db 71 GCTTGTCTCCCGGAGACGCTATCTCATCTCTTCTTTCGATTAACCGTGGCGGAG 130
QY 120 AGTCAGGCGCGCGCTGCGGACGACGAGCGCGGCGGCGGCGGCGGCGGCGGCGGAG 179
Db 131 AGTCAGGCGCGCGCTGCGGACGACGAGCGCGGCGGCGGCGGCGGCGGCGGCGGAG 190
QY 180 ATGTCCAGCATGATCCCGAATATGATTAATTAATTAATTAATTAATTAATTAATTAAT 239
Db 191 ATGTCCAGCATGATCCCGAATATGATTAATTAATTAATTAATTAATTAATTAATTAAT 250
QY 240 GGGGTGGAAGTCTTCT 299
Db 251 GGGGTGGAAGTCTTCT 310
QY 300 ATCAGCACAATTTGGTGGATTTCAAAATTAAGACTATATAGAGTTAGACGGAAAAACATC 359
Db 311 ATCAGCACAATTTGGTGGATTTCAAAATTAAGACTATATAGAGTTAGACGGAAAAACATC 370
QY 360 AAGCTCAATTAAGAGTCTCTCAATTAATTTAAACAGTGGCGGACGGAATATGATCTTAT 419
Db 371 AAGCTCAATTA----- 382
QY 420 GCCAGTGAATATGTCAACAATTTGTTGTAAGGAACAATGTATCTGACCAAAAGAA 479
Db 383 ----- 382
QY 480 CTAGTACACTACACAACAGCAAGAAATTTGCTGATTCCTTGGAAATTCGTTTTGGAA 539
Db 383 ----- GATTTGCTGATTCCTTGGAAATTCGTTTTGGAA 418
QY 540 ACCAGTCTAAGATGACAGCAATGTGAAACAGTCTTTCATGACAGTGGACCTGAGATT 599
Db 419 ACCAGTCTAAGATGACAGCAATGTGAAACAGTCTTTCATGACAGTGGACCTGAGATT 478
QY 600 AAAAGCAATGGTCCGAGCAACAGCTGTGTGTGTGAGAGTCAATGTTAAATT 659
Db 479 AAAAGCAATGGTCCGAGCAACAGCTGTGTGTGTGAGAGTCAATGTTAAATT 538
QY 660 CAGAGCACTCCAGTCAACAGTCAAGTGTGAGTGTGCTAAATTTGGCTTCATCTCTT 719
Db 539 CAGAGCACTCCAGTCAACAGTCAAGTGTGAGTGTGCTAAATTTGGCTTCATCTCTT 598
QY 720 TCTCAGCAAGATGATTTGCAATCTGAACCAAGTGAAGAAAAACAATTTGCTGAATTTG 779
Db 599 TCTCAGCAAGATGATTTGCAATCTGAACCAAGTGAAGAAAAACAATTTGCTGAATTTG 658

OY	780	ACGATATGAGTGGACATACAAACAGATATTCCTACCGTCACCAAAAGTCAGAGATTGTAA	839
Db	659	ACTGATATGAGTGGACATACAAACAGATATTCCTACCGTCACCAAAAGTCAGAGATTGTAA	718
OY	840	ATGTCATATCTGACTTTTATTTTATTTTCCTTGACTCAAGACGCTAACATCTATTTTCAG	899
Db	719	ATGTCATATCTGACTTTTATTTTATTTTCCTTGACTCAAGACGCTAACATCTATTTTCAG	778
OY	900	AACGTGTTTAAACCTTTGTGTGCTGGTTATTAATAATATGTGTATATCTTGTGCTTT	959
Db	779	AACGTGTTTAAACCTTTGTGTGCTGGTTATTAATAATATGTGTATATCTTGTGCTTT	838
OY	960	CCTGATACAGACGCTTTCCCGTGGTGGTTAGATATATTTGTGTTGATGTTATAT	1019
Db	839	CCTGATACAGACGCTTTCCCGTGGTGGTTAGATATATTTGTGTTGATGTTATAT	898
OY	1020	GGCATGTTAGATGTCAGGTTAGTCTCTTGAAGATGAAGTTCACCCATTTGTATCANA	1079
Db	899	GGCATGTTAGATGTCAGGTTAGTCTCTTGAAGATGAAGTTCACCCATTTGTATCANA	958
OY	1080	CAGCAACAAGCAGTGTCTGTCACTTTCATGCAATAAGTTTACTGAGATGTATATGTAA	1139
Db	959	CAGCAACAAGCAGTGTCTGTCACTTTCATGCAATAAGTTTACTGAGATGTATATGTAA	1018
OY	1140	ATTCATATTTGCAATGCTTCCCTGTGAGATTAATAATGGAAGATTACACTATCGATTA	1199
Db	1019	ATTCATATTTGCAATGCTTCCCTGTGAGATTAATAATGGAAGATTACACTATCGATTA	1078
OY	1200	ATAGTTTCTTCATCTGCTCATATATTTGTGGCTGCAGATATTTGATTTGTTGCACA	1259
Db	1079	ATAGTTTCTTCATCTGCTCATATATTTGTGGCTGCAGATATTTGATTTGTTGCACA	1138
OY	1260	CTATGTAAACAAACAACGTAGATATGTTTAATAATATTTGTAATTTGGAAGTAATAAA	1319
Db	1139	CTATGTAAACAAACAACGTAGATATGTTTAATAATATTTGTAATTTGGAAGTAATAAA	1198
OY	1320	AAAAAAAAAAAA 1329	
Db	1199	AAAAAAAAAAAA 1208	

RESULT 5				
MMYPT1				
LOCUS	1428 bp	mRNA	linear	ROD 12-SEP-1993
DEFINITION	Mouse mRNA for ras-related ypt1 protein.			
ACCESSION	Y00094			
VERSION	Y00094.1	GI:55456		
KEYWORDS	ras-related protein; ypt1 gene.			
SOURCE	Mus musculus.			
ORGANISM	Mus musculus.			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 1428)			
AUTHORS	Gallwitz,D.			
TITLE	Direct Submission			
JOURNAL	Submitted (13-NOV-1987) Gallwitz D., Max-Planck-Institut fuer			
	biophysikalische Chemie, Abt. Molekulare Genetik, Postfach 2841,			
	D-3400 Goettingen, FRG			
	2 (bases 1 to 1428)			
REFERENCE	Hadravsky,H., Disela,C., Wagner,P. and Gallwitz,D.			
AUTHORS	The ras-related ypt1 protein is an ubiquitous eukaryotic protein:			
TITLE	Isolation and sequence analysis of mouse cDNA clones highly			
	homologous to the yeast ypt1 gene			
JOURNAL	EMBO J. 6 (1987) In press			
FEATURES	Location/Qualifiers			
SOURCE	1..1428			

CDS

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="F9 and C3H10T1/2"
130. 747
/note="ppl protein (AA 1-205)"
/codon_start=1
/protein_id="CAA68284.1"

```

[illegible]

OY	1075	TCAAACACACACAGAGAGTGTGCACGTCATGCATAAAGTTTGTAGATGTTAAT	1134
Db	1020	TCACACACACACACAGAGTGTGCACGTTTCCACGCATAAAGTTTGTAGACGTTAAT	1079
OY	1135	GTAAGATCTGATTTGCTAGTCTTCCCTGTGAGATTAAATGAAAGATTACACTATCT	1194
Db	1080	GTAAGATCTGATTTCTAGTCTTCCCTGTGAGATTAAAGAAAGATTACACTATCT	1139
OY	1195	GATTAAATGTTCTTCATACCTCGATATTAATTGGCGTGGAGATATTGAAATTGTT	1254
Db	1140	GATTAAATGTTCTTCATACCTCGATATTAATTGTGGCTGCAGATATTGTAATTTGTT	1199
OY	1255	GCACACTATGTAAACAAACAACCTGAGATATGTTTAATAATATTGTACTTAATGGAACT	1314
Db	1200	GCACACTATGTAAACAA--AACTGAAGATATGTTTAATAATATTGTACTTAATGGAACT	1256
OY	1315	AAAAAAAAA 1323	
Db	1257	AATATCAAA 1265	

RESULT 7	AC123143	LOCUS	DEFINITION
	AC123143	151988 bp	DNA linear HTG 17-JUL-2007
		Rattus norvegicus clone CH230-247B24,	*** SEQUENCING IN PROGRESS
		*** 50 unordered pieces.	

ACCESSION	AC123143
VERSION	AC123143..2
KEYWORDS	HTG, HTGS_PHASE1.
SOURCE	Norway IdL.
ORGANISM	Rattus norvegicus

REFERENCE
AUTHORS
1 (bases 1 to 151988)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,

Albarrado, S.L., Amarantunge, H.C., Are, J.R., Ayele, A.K., Banks, T.,
Barbarea, J., Benton, J., Blinage, K., Blankenburg, K., Bonini, D.,
Bouck, J., Bowle, S., Brierley, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Devila, M.L., Davis, C., Davy-Carroll, L., Deedrich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabalis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorelli, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Homsli, F., Howard, S., Huber, J., Hult, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S.,
Karlsone, E., Kelly, S., Khan, U., King, L., Korval, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Ilean, C., Liu, J., Liu, W., Louisgeed, H.,
Lorado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., McWhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Mosier, M., Neal, D., Newton, J.J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokweto, S., Oguh, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peterson, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Schierer, S., Scott, G., Shen, H., Shoostalt, N., Sisson, I.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Swatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 151986)
AUTHORS Morley, R.C.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Human Genome Sequencing Center, Department of Genetics, Baylor College of Medicine, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 151988)

JOURNAL Submitted (17-JUL-2002) Human Genome Sequencing Center, Department

COMMENT On Jul 4, 2002 this sequence version replaced gl:21239895.

Center code: BCM

----- Project Information

----- Summary Statistics -----

Assembly program: Phrap; version 0.990329

Consensus quality: 115635 bases at least Q20

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

* is not known and their order in this sequence record is

* This record will be updated with the finished sequence

*	1	1041: contig of 1041 bp in length
---	---	-----------------------------------

*	2642	2741: gap of unknown length
---	------	-----------------------------

Accession	Contig	Length (bp)
3932	contig of 1685 bp	1685

*	6733	6832: gap of unknown length
---	------	-----------------------------

Accession	Gene	Length (bp)
8290	contig of 1121 bp in length	1121
9410	contig of 1121 bp in length	1121

```

* 10655 2000:1: connect: connection refused
* 10754: gap of unknown length

```

Accession	Gene	Length (bp)
11948	gap of unknown length	13492
*	contig of 1545 bp in length	

```

14688      *
14689      14787: gap of unknown length

```

	gap or unknown length
10132	10231:
16252	18133: contig of 1882 bp in length

*	10234	15003: conflict of 1032 by in length
*	19866	19965: gap of unknown length

```
* 21210 21309: gap of unknown length
* 21310 21382: contig of 1873 bp in length
* 21313 23282: gap of unknown length
* 21320 24319: contig of 1037 bp in length
* 21323 24419: gap of unknown length
* 24420 26547: contig of 2128 bp in length
* 26548 26647: gap of unknown length
* 26648 28422: contig of 1775 bp in length
* 28423 28532: gap of unknown length
* 28533 30850: contig of 2328 bp in length
* 30851 30950: gap of unknown length
* 30951 32773: contig of 1823 bp in length
* 32774 32873: gap of unknown length
* 32874 35270: contig of 2397 bp in length
* 35271 35370: gap of unknown length
* 35371 37182: contig of 1812 bp in length
* 37183 37282: gap of unknown length
* 37283 38841: contig of 1559 bp in length
* 38842 38941: gap of unknown length
* 38942 41828: contig of 2887 bp in length
* 41829 41929: gap of unknown length
* 41929 44919: contig of 2991 bp in length
* 44920 45019: gap of unknown length
* 45020 48137: contig of 3118 bp in length
* 48138 48237: gap of unknown length
* 48238 50153: contig of 1916 bp in length
* 50154 50253: gap of unknown length
* 50254 53223: contig of 2970 bp in length
* 53224 53283: gap of unknown length
* 53284 58269: contig of 4946 bp in length
* 58270 58369: gap of unknown length
* 58370 61783: contig of 3414 bp in length
* 61784 61883: gap of unknown length
* 61884 65284: contig of 3401 bp in length
* 65285 65385: gap of unknown length
* 65386 68318: contig of 2934 bp in length
* 68319 68419: gap of unknown length
* 68419 72259: contig of 3841 bp in length
* 72260 72359: gap of unknown length
* 72360 74905: contig of 2546 bp in length
* 74906 75005: gap of unknown length
* 75006 78675: contig of 3670 bp in length
* 78676 78775: gap of unknown length
* 78776 81774: contig of 2999 bp in length
* 81775 81874: gap of unknown length
* 81875 86361: contig of 4487 bp in length
* 86362 86461: gap of unknown length
* 86462 90372: contig of 3911 bp in length
* 90373 90472: gap of unknown length
* 90473 94086: contig of 3614 bp in length
* 94087 94186: gap of unknown length
* 94187 97642: contig of 3456 bp in length
* 97643 97742: gap of unknown length
* 97743 104964: contig of 7222 bp in length
* 104965 105064: gap of unknown length
* 105065 111556: contig of 6492 bp in length
* 111557 111657: gap of unknown length
* 111658 115627: contig of 3971 bp in length
* 115628 115727: gap of unknown length
* 115728 122141: contig of 6414 bp in length
* 122142 122411: gap of unknown length
* 122412 127872: contig of 5631 bp in length
* 127873 127972: gap of unknown length
* 127973 133317: contig of 5345 bp in length
* 133318 133417: gap of unknown length
* 133418 139240: contig of 5823 bp in length
* 139241 143394: gap of unknown length
* 143395 143494: contig of 4034 bp in length
* 143495 151988: contig of 8494 bp in length.
Location/Qualifiers
1. .151988
/organism="Rattus norvegicus"
/db.xref="taxon:10116"
/clone="CH230-247B24"
BASE COUNT 44145 a 28740 c 28161 g 43504 t 7438 others
ORIGIN
Query Match 58.4%; Score 820.6; DB 2; Length 151988;
Best Local Similarity 78.8%; Pred. No. 5,8e-123;
Matches 1196; Conservative 0; Mismatches 114; Indels 207; Gaps 10;
1 AAGCATATGCTAGCT-GCGCGGCGGCTGATTTGTCTAGGAGGAGGAGTGGAGAGA 59
Db 75930 AAGCATATGCTAGCTGCTGAGGCGCTGATTTGTATTAAGCTCGCTG-CGGTTGA 75988
OY 60 CGTTGCTGCTCCGGAACGCTATTCATCTCTTCTTGTGATTAACCGTGGAG 119
Db 75989 CGTTGCTGCTGGAACGCTATTCATCTCTTCTTGTGATTAACCGTGGAG 76048
OY 120 AATCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 179
Db 76049 GCTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 76105
OY 180 ATGTCAGCATGAAATCCGAAATGATATTTATCAATGATCTGATGCGACATCA 239
Db 76106 ATGTCAGCATGAAATCCGAAATGATATTTATCAATGATCTGATGCGACATCT 76165
OY 240 GGGGTTGGAAGCTCTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 299
Db 76166 GGGGTTGGAAGCTCTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 76225
OY 300 ATCAGCAATTTGTTGATTTCAAAATAGAACTATAGATTAGACGGGAAAAATTC 359
Db 76226 ATCAGCAATTTGTTGATTTCAAAATAGAACTATAGATTAGACGGGAAAAATTC 76285
OY 360 AAGCTTCATAATV----- 371
Db 76286 AAGCTTCATATATGGGACACAGACAGAGAAAGATTTGCAACATCACTCCAGTTAT 76345
OY 372 -----GAGTCTTCAT 383
Db 76346 TACAGAGAACCATTTATCATATGATGATGATGATGATGATGATGATGATGATGAT 76405
OY 384 AATGTTAAAGAGTGGCTGACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 443
Db 76406 AATGTTAAAGAGTGGCTGACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 76465
OY 444 TTGGTAGGGAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 503
Db 76466 TTGGTAGGGAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 76521
OY 504 GAATTTGCTATTCCTTGGAAATTCGTTTGGAAATTCGTTTGGAAATTCGTTTGGAAATTCGTTT 563
Db 76522 GAATTTGCTATTCCTTGGAAATTCGTTTGGAAATTCGTTTGGAAATTCGTTTGGAAATTCGTTT 76581
OY 564 GTAGAACAGCTTTCTATGACATGCGACGCTGAGATTTAAAGAGATGGTCCCGAGCA 623
Db 76582 GTAGAACAGCTTTCTATGACATGCGACGCTGAGATTTAAAGAGATGGTCCCGAGCA 76641
OY 624 ACAAGTGTGTGCTGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGAGTCA 683
Db 76642 ACAAGTGTGTGCTGAGAGTCCATGTTAAATTCAGAGCACTCCAGTCAAGAGTCA 76701
OY 684 GGTGAGAGTGTGCTGTAATAATTTGCTTCATCCCTTTCTCAGACAAATGATTTGCAATTC 743
Db 76702 GGTGAGAGTGTGCTGTAATAATTTGCTTCATCCCTTTCTCAGACAAATGATTTGCAATTC 76761
OY 744 TGAACCCAGTGAATAAATTCGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 803
Db 76762 TGAACCCAGTGAATAAATTCGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 76821
OY 804 GATTTCTACGCTCCACAAAGGTCAGAGATGTTAAAGTCATATCTGAC-TTTT 862
Db 76822 GATTTCTACGCTTTTCAC-AAAGTCAGAGATGTTAAAGTCATATCTGAC-TTTT 76880
```

QY 863 TATTCCTGACTCAAGACAGTCACTTTCAGACTGTTTAAACCTTGTGTC 922
 Db 76881 TATTCCTGACTCAAGACAGTCACTTTCAGACTGTTTAAACCTTGTGTC 76940
 QY 923 TGGTTTAAATAATATGCTGTAATCCCTGTTGCTTCTGATACCAACCTGTTCCCGT 982
 Db 76941 TGGTTTAAATAATATGCTGTAATCCCTGTTGCTTCTGATACCAACCTGTTCCCGT 77000
 QY 983 GGTGCTAGAAATATATTTGTTGATGTTTATATATGTCAGTGT----- 1027
 Db 77001 GGTGCTAGAAATATATTTGTTGATGTTTATATATGTCAGTGTGTTG 77060
 QY 1028 ----- 1027
 Db 77061 GTTGGAGATTAGCTAGTGTAGAGCGCTTGCTAGCAAGCGCCGTGGTTCG 77120
 QY 1028 -----TAGATGTCAGTTTATAGTCTTCTGAAAGATGAAGTTACAGC 1066
 Db 77121 TCCCGAGCTCCGAAAAAAATAGATGTTGGTTAGTCTTCTGAAAGATGAAGTTACAGC 77180
 QY 1067 ATTTGATCAAAACAGACAGCACTGCTGTCACCTTCCATGCAATGAAGTTAGTGA 1126
 Db 77181 ATTTGATCAAAATGACACACACACAGCTGCTGTCACCTTCCATGCAATGAAGTTAGTGA 77240
 QY 1127 TGTATATGTAAGATGCTGATTTGCTAGTCTTCTGTAAGATTAATAATGGAAGATTA 1186
 Db 77241 CATATATGTAAGATGCTGATTTGCTAGTCTTCTGTAAGATTAATAATGGAAGATTA 77299
 QY 1187 CACATATGTAAGATGCTGATTTGCTAGTCTTCTGTAAGATTAATAATGGAAGATTA 1246
 Db 77300 CACATATGTAAGATGCTGATTTGCTAGTCTTCTGTAAGATTAATAATGGAAGATTA 77359
 QY 1247 AATTTGTCACACATGATTAACAACAACAGTGAAGATGTTTAAATATGTTACTTA 1306
 Db 77360 AATTTGTCACACATGATTAACAACAACAGTGAAGATGTTTAAATATGTTACTTA 77416
 QY 1307 TTGGAAGTAAAAAAA 1323
 Db 77417 TCGGAAGTAAATATCAAA 77433

RESULT 8
 AC007318/c 204230 bp DNA linear PRI 07-NOV-2001
 LOCUS AC007318 Homo sapiens BAC clone RP11-420C9 from 2, complete sequence.
 DEFINITION AC007318.4 GI:13435282
 ACCESSION AC007318.4
 VERSION AC007318.4
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 204230)
 TITLE Toward a complete human genome sequence
 AUTHORS Sulston, J.E. and Waterston, R.
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 204230)
 TITLE The sequence of Homo sapiens BAC clone RP11-420C9
 AUTHORS Wolian, A., Hawkins, M., Kalicki, J., Harting, R., and Lehnert, L.
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 204230)
 TITLE Direct Submission
 AUTHORS Submitted (16-APR-1999) Genome Sequencing Center, Washington
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 4 (bases 1 to 204230)
 TITLE Direct Submission
 AUTHORS Submitted (23-MAR-2001) Genome Sequencing Center, Washington
 JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE MO 63108, USA
 AUTHORS 5 (bases 1 to 204230)
 TITLE Waterston, R.
 JOURNAL Direct Submission
 REFERENCE Submitted (09-AUG-2001) Department of Genetics, Washington
 TITLE University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 AUTHORS 6 (bases 1 to 204230)
 JOURNAL Waterston, R.
 TITLE Direct Submission
 COMMENT Submitted (07-NOV-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 23, 2001 this sequence version replaced gi:7630984.

 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.wustl.edu

 Center project name: H_NH0420C09

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e. phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
 Taleno, M., Catanesi, J.J., de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (http://www.resgen.com) or Pter de Jong
 and coworkers at the Roswell Park Cancer Institute
 (http://bacpac.med.buffalo.edu)
 VECTOR: pBAC3.6
 NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-350H15; the clone sequenced
 to the right is AC034121. Actual start of this clone is at base
 position 1 of RP11-420C9; actual end is at base position 204230 of
 RP11-420C9.

The sequence RP11-420C9 from base position 44230 to 46506 consists
 of PCR product from clone DNA. Base position 44700 to 44774
 represents sequence of a single plasmid subclone derived from PCR
 amplification of clone DNA.

FEATURES
 source
 1. 204230
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-420C9"
 /clone_id="RP11-11"
 179. 486
 /rft_family="Alu"
 890. 1082
 /rft_family="MER1-type"
 1498. 1789
 repeat_region
 repeat_region
 repeat_region
 repeat_region

	misc_feature	6663..6949	/note=*similar to EST T07351 (NID:g318500) "
	misc_feature	6861..7737	/note=*similar to EST A0124391 (NID:g1094107) "
	misc_feature	7006..7434	/note=*similar to EST AB825455 (NID:g2896766) oee4f08.s1"
	misc_feature	7130..7930	/note=*similar to EST BF242402 (NID:g11156329) "
	repeat_region	7171..7467	
	misc_feature	7202..7396	/note=*_family="Alu"
	misc_feature	7501..8188	EST AV703078 (NID:g10719408) "
	misc_feature	7650..7920	EST BE279740 (NID:g9154735) "
	misc_feature	7790..8251	EST AV703078 (NID:g10719408) "
	misc_feature	7945..8533	EST AA255857 (NID:g1891513) zsz7rh08.r1"
	misc_feature	8047..8633	/note=*similar to EST AW601747 (NID:g9724069) "
	misc_feature	8065..8647	/note=*similar to EST AW550283 (NID:g7195711) "
	misc_feature		/note=*similar to EST AU49130 (NID:g11010651) "
	Query Match	58.2%; Score 817.2; Db 9;	Length 204230;
	Best Local Similarity	99.6%; Ped.	No. 2.1e-122;
	Matches 819;	Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	502	AGCAATTTGCTAATTCCCTTGGAATTCGGTTTTTGGAAACCGTCTTAAGAATGCAMCA	561
Db	10336	AGGAATTTGCTAATTCCTTGGAATTCGGTTTTTGGAAACCGTCTTAAGAATGCAMCA	10277
QY	562	ATGAGAACAGTCCTTTCATGACANTGGCAGGTGAGATTTAAAAAGCGAATGGCTCCGGAG	621
Db	10276	ATGAGAACAGTCCTTTCATGACANTGGCAGGTGAGATTTAAAAAGCGAATGGCTCCGGAG	10217
QY	622	CACAAGCTGGTGGTGGTAGAAGCTCCATGTTAAAATTAGAGCACGCCAGTCAAGCACT	681
Db	10216	CACAAGCTGGTGGTGGTAGAAGCTCCATGTTAAAATTAGAGCACGCCAGTCAAGCACT	10157
QY	682	CAGGTGGAGGTGGTGTAAAAATTTGCCCATCCTTTTCTACAGCAATGAATTTGCAA	741
Db	10156	CAGGTGGAGGTGGTGTAAAAATTTGCCCATCCTTTTCTACAGCAATGAATTTGCAA	10097
QY	742	TCTAACACCAGTGAATAAAACAAATAGCCTGATGTGACTGATATGAGTGCACATCAA	801
Db	10096	TCTAACACCAGTGAATAAAACAAATAGCCTGATGTGACTGATATGAGTGCACATCAA	10037
QY	802	CAGATTTTACGCTCCACAAGAGTCAGAAATGTAAATGGTCAATACAGACTTTTTT	861
Db	10036	CAGATTTTACGCTCCACAAGAGTCAGAAATGTAAATGGTCAATACAGACTTTTTT	9977
QY	862	TTATTCCTTGGACTCAGACAGCTAATCTATTTTGAAACGCTTTTAAACCTTTGTGTG	921
Db	9976	TTATTCCTTGGACTCAGACAGCTAATCTATTTTGAAACGCTTTTAAACCTTTGTGTG	9917
QY	922	CTGGTTTAAATAAATAATGTGTGAATCCCTTTGCTCTTTCCTGATATACAGACGTTTCCG	981
Db	9916	CTGGTTTAAATAAATAATGTGTGAATCCCTTTGCTCTTTCCTGATATACAGACGTTTCCG	9857
QY	982	TGCTTGGTAGAATATATTTTGTTTGTAAGTTTAAATTTGGCATGTTTAGATGTCAAGTTT	1041
Db	9856	TGCTTGGTAGAATATATTTTGTTTGTAAGTTTAAATTTGGCATGTTTAGATGTCAAGTTT	9797
QY	1042	AGTCTTCTAAGATTAAGTTCACCCATTTTGTATCAAAACAGCAACACAGATGTCTGCAC	1101
Db	9796	AGTCTTCTAAGATTAAGTTCACCCATTTTGTATCAAAACAGCAACACAGATGTCTGCAC	9737
QY	1102	TTTCCATGATTAATTTAGTGAATGTAATAGTAAGATCTGAATTTGGTACTGCTTCCCT	1161
Db	9736	TTTCCATGATTAATTTAGTGAATGTAATAGTAAGATCTGAATTTGGTACTGCTTCCCT	9677

[illegible]

Db	17439	ATGTAGAAACATCTTTTATCATGACGATGGCAGCTGATTAATAAACGAATGGCTCTGGAG	177498
Qy	622	CACACGCTGGTGGTCTGAGAGTCCAAATGTTAAATTCAGAGCAGTCCAGCAGT	681
Db	177499	CTACAGCGTGGTGGCCGAGAGTCCATGTTAAATTCAGAGCAGTCCAGCAGT	177558
Qy	682	CAGGTGGAGGTGGTGGCAAAATTTGGCTCCATCTCTTTCACAGCAATGAATTTGCA	741
Db	177559	CAGGTGGAGGTGGTGGCAAAATTTGGCTCCATCTCTTTCACAGCAATGAATTTGCA	177618
Qy	742	TCTGAAACCAGTAAAAAAACAAATTTGCCGAATTTGACTGTATGTAGTCGACATCAA	801
Db	177619	TCTGAAACCAGTAAAAAAACAAATTTGCCGAATTTGACTGTATGTAGTCGACATCAA	177678
Qy	802	CAGATCTTTACCGTCTCCACAAGGTCAGAGATTTGTAAATGCTCAATACGAC--TTT	860
Db	177679	CAGATCTTTACCGTCTCCACAAGGTCAGAGATTTGTAAATGCTCAATACGAC--TTT	177737
Qy	861	TTTATTTCCCTTGACGCAAGACAGCACTTCAATTTTGACAGCTGTTTAAACCTTTGGT	920
Db	177738	TTTATTTCCCTTGACGCAAGACAGCACTTCAATTTTGACAGCTGTTTAAACCTTTGGT	177797
Qy	921	GCTGGTTTATTAATATATGTGTATATCTTGTGCTTTCCCTGATACAGACGTGTTCC	980
Db	177798	GCTGGTTTATTAATATATGTGTATATCTTGTGCTTTCCCTGATACAGACGTGTTCC	177857
Qy	981	GTGGTGGTTAGAAATATATTTTGTGATGTGTTATATATGGCAATGTTAGATGTCAGTT	1040
Db	177858	GTGGTGGTTAGAAATATATTTTGTGATGTGTTATATATGGCAATGTTAGATGTTAGTT	177917
Qy	1041	TAGTCTTCTGAAAGTGAAGTCCAGCCATTTGTATGCAAAAGCAAGCAGTGTCTGCA	1100
Db	177918	TAGTCTTCTGAAAGTGAAGTCCAGCCATTTGTATGCAAAAGCAAGCAGTGTCTGCA	177977
Qy	1101	CTTTCATGATCAATTAAGTTAGTGAATGTTATATGTAAGATCTATTTGCTAGTCTTCC	1160
Db	177978	CTTTCATGATCAATTAAGTTAGTGAATGTTATATGTAAGATCTATTTGCTAGTCTTCC	178037
Qy	1161	TTGTAGACTTATTAATGGAAGATTAACATCTTGATTAATAGTTTCTTCACTACTGCA	1220
Db	178038	TTGTAGACTTATTAATGGAAGATTAACATCTTGATTAATAGTTTCTTCACTACTGCA	178097
Qy	1221	TATATTTGTGGCGCACAATTTGTATTTGTGGCACACATGTAACAAAACACTGAA	1280
Db	178098	TATATTTGTGGCGCACAATTTGTATTTGTGGCACACATGTAACAAA--AACTGAA	178154
Qy	1281	GATATGTTTAAATATATTTGACTTATTTGGAAGTAAAAAAA	1323
Db	178155	GATATGTTTAAATATATTTGACTTATTTGGAAGTAAAAAAA	178197
RESULT 10			
AL512349/c			
LOCUS	AL512349	116614 bp	DNA linear HTG 10-JUL-2001
DEFINITION	Homo sapiens chromosome 6 clone RP11-223J2, *** SEQUENCING IN		
ACCESSION	AL512349		
VERSION	AL512349.7	GI:13121473	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_CANCELLED.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Bates, K.		
JOURNAL	Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,		
COMMENT	CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk		
	On Feb 23, 2001 this sequence version replaced gi:12956981.		
	----- Genome Center		
	Center: Sanger Centre		

Accession AL049545
Version AL049545.6 GI:5002650
Keywords RFG: 60S Ribosomal Protein L7; RAB1; RAS; RPL7.
Source Homo sapiens.
Organism Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Reference 1 (bases 1 to 104081)
Author Smith, S.
Title Direct Submission
Journal Submitted (06-JUL-1999) Sanger Centre, Hinxton, Cambridgeshire, CE10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
Comment requests: clonerequest@sanger.ac.uk
On Jun 7, 1999 this sequence version replaced gi:4835284.
During difference assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL.
This sequence is the entire insert of clone 263J7. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>
263J7 is from the library RPCI1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:PCYPAC2>.
FEATURES
source
1.104081
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="q14.3-15"
/clone="RP1-263J7"
/clone_1lb="RPCI-1"
1..4880
/note="LIP5 repeat: matches -237..5623 of consensus"
4895..5820
/note="LIP5 repeat: matches -1405..-405 of consensus"
5861..6011
/note="MLTJ repeat: matches 36..174 of consensus"
6032..6086
/note="L1 repeat: matches 3626..3680 of consensus"
6740..8505
/note="LIP12 repeat: matches 4282..6165 of consensus"
8512..9328
/note="LIP repeat: matches 3049..3866 of consensus"
9329..9626
/note="AluYb8 repeat: matches 1..317 of consensus"
9627..10966
/note="LIP repeat: matches 1712..3049 of consensus"
complement(11030..11749)
/gene="dJ263J7.1"
/gene="dJ263J7.1"
complement(11030..11749)
/gene="dJ263J7.1"
/note="dJ263J7.1 (RPL7 (60S Ribosomal Protein L7) pseudogene): match: cDNAs: Em:X52967 Em:X57959 Em:LI6558 Em:X57958 Em:M85235 Em:M17422 Em:M29016 Em:X57961 Em:X57960 Em:U63785; match: ESTs: Em:AL047212 Em:LI683841 Em:AI609411 Em:AI620251 Em:AI678713 Em:AA853290 Em:AI188448 Em:AI246632 Em:F28697.1 Em:AA640941 Em:AA147085 Em:AI354542 Em:AA648558 Em:AA172041

Em:AA736700 Em:AA181027 Em:AA617963 Em:AA129663 Em:AI385514 Em:AA58684 Em:AA148321 Em:AA147910 Em:AA065193 Em:AA223084 Em:AA065240 Em:AA147610 Em:AI498784 Em:AA313779 Em:AA626741 Em:AA397390 Em:AA639843 Em:AA598466 Em:AA587875 Em:AA314579 Em:AA081412 Em:AA838601 Em:AA129729 Em:AA236303 Em:AA580325 Em:AA147597 Em:AA313320 Em:AA516160 Em:AA311117 Em:AA562608 Em:AA635115 Em:AI367854 Em:AA308128 Em:AA846920 Em:AA608110 Em:AA745100 Em:AA607384 Em:AA507813 Em:AA314838 Em:AA736534 Em:AA591026 Em:AA008632 Em:W77465 Em:AA648073 Em:AA648202 Em:AA552160 Em:AA715370 Em:DS1697 Em:DS1979 Em:DS5053 Em:AA760175 Em:AA856854 Em:N98334 Em:AA649499 Em:AA879164 Em:AA332842 Em:AA886868 Em:AA173529 Em:AA091272 Em:AA093750; match: proteins: Sw:PI8124 Sw:PI4148 Sw:PO5426 Sw:Q42208 Sw:O01802 Sw:O60143 Sw:P25457 Sw:PI1874 Sw:P32100 Sw:P05737 Sw:Q12213 Tr:O14371 Tr:Q08953"
/pseudo
/codon_start=-1
/evidence-not_experimental
11307..11789
/note="match: GSS A0801091 B54651 A0028779"
11790..12860
/note="LIP13 repeat: matches 647..1730 of consensus"
12908..13190
/note="AluJo repeat: matches 3..302 of consensus"
13341..13560
/note="HAI1 repeat: matches 1483..1721 of consensus"
13346..13347
/note="3 subclones read TCTT, 2 read TCT"
13594..15880
/note="LIPB2 repeat: matches 3950..6154 of consensus"
15881..16608
/note="LIPB2 repeat: matches 5413..6144 of consensus"
16609..19090
/note="LIPB4 repeat: matches 1394..3950 of consensus"
19096..19323
/note="LIPB2 repeat: matches 1394..3950 of consensus"
19472..23244
/note="LIP3 repeat: matches 3..231 of consensus"
23240..23462
/note="LIP repeat: matches 638..4389 of consensus"
23577..29885
/note="LIPBa repeat: matches -1549..-1330 of consensus"
29886..29943
/note="LIPB3 repeat: matches -218..6130 of consensus"
29944..30234
/note="LIPAS repeat: matches 6081..6139 of consensus"
30235..30912
/note="LIPB3 repeat: matches -508..-218 of consensus"
complement(30558..30912)
/note="match: EST AA528504"
31385..31601
/note="MIR repeat: matches 29..256 of consensus"
31721..32002
/note="LIP repeat: matches 5862..6145 of consensus"
32279..32583
/note="AluSx repeat: matches 1..308 of consensus"
33356..33592
/note="MIR repeat: matches 8..247 of consensus"
33779..37507
/note="LIM42 repeat: matches 2533..6308 of consensus"
37508..37811
/note="AluY repeat: matches 3..305 of consensus"
37812..37985
/note="LIM42 repeat: matches 2362..2533 of consensus"
37976..41630
/note="LIM1 repeat: matches -1390..2698 of consensus"
41878..41924
/note="MER5B repeat: matches 127..173 of consensus"
42325..42418
/note="MAD1 repeat: matches 1..79 of consensus"
42451..42853
/note="MLTIC repeat: matches 23..466 of consensus"

Qy 1307 TTGGAGTAAA 1366
 Db 94457 CAATGGATCTAATTAAACCAAGAGCTTCTGTCAGAAATTAACCTAGACACAGAGTG 94398
 Qy 1367 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1394
 Db 94397 AATAGACACACACAAATGAGAGAAA 94370

RESULT 12
 AC067945/c 162471 bp DNA linear PRI 07-NOV-2001
 LOCUS AC067945 Homo sapiens BAC clone RP11-629B4 from 2, complete sequence.
 DEFINITION AC067945 GI:13786486
 VERSION HTG.
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 162471)
 TITLE Tulsion,J.E. and Waterston,R.
 JOURNAL Toward a complete human genome sequence
 MEDLINE Genome Res. 8 (11), 1097-1108 (1998)
 PUBMED 99063792
 9847074
 2 (bases 1 to 162471)
 AUTHORS Isak,A., Elliott,G., Doeber,A., Abbott,A., Hawkins,M. and Falk,A.
 TITLE The sequence of Homo sapiens BAC clone Rp11-629B4
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 162471)
 AUTHORS Waterston,R.H.
 JOURNAL Direct Submission
 Submitted (27-APR-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 162471)
 AUTHORS Waterston,R.H.
 JOURNAL Direct Submission
 Submitted (25-APR-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE
 AUTHORS 5 (bases 1 to 162471)
 TITLE Waterston,R.
 JOURNAL Direct Submission
 Submitted (09-AUG-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 6 (bases 1 to 162471)
 AUTHORS Waterston,R.
 JOURNAL Direct Submission
 Submitted (07-NOV-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Apr 25, 2001 this sequence version replaced gl:13518285.

COMMENT
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0629B04

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 The RPCI-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
 Tatenno,M., Cateneze,J.J. and de Jong,P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (http://www.resgen.com) or Pletier de Jong
 and coworkers at the Roswell Park Cancer Institute
 (http://bacpac.med.buffalo.edu)
 VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-45219, 200 bp overlap; the
 clone sequenced to the right is RP11-31799. Actual start of this
 clone is at base position 195 of RP11-629B4; actual end is at base
 position 162471 of RP11-629B4.

FEATURES

Source

1. 162471
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-629B4"
 /clone_11b="RPCI-11"
 777. 1067
 /rpt_family="Alu"
 1352. 1380
 /rpt_family="AT_rich"
 1887. 2071
 /rpt_family="MIR"
 3209. 3257
 /rpt_family="MIR"
 4158. 4475
 /rpt_family="L1"
 4476. 4775
 /rpt_family="Alu"
 4776. 5044
 /rpt_family="L1"
 5108. 5242
 /rpt_family="Alu"
 5428. 5465
 /rpt_family="AT_rich"
 5617. 5910
 /rpt_family="Alu"
 5997. 6133
 /rpt_family="Alu"
 6135. 6178
 /rpt_family="Alu"
 6213. 6498
 /rpt_family="Alu"
 9226. 9268
 /rpt_family="AT_rich"
 9573. 9600
 /rpt_family="AT_rich"
 10528. 10562
 /rpt_family="AT_rich"
 11440. 11467
 /rpt_family="(TTTGG)n"
 12297. 12325
 /rpt_family="AT_rich"
 12472. 12518
 /rpt_family="(TATANG)n"
 14656. 14747
 /rpt_family="CRI"
 16030. 16089

repeatL_region	/rpt_family="AT-rich"	17221..17290	
repeatL_region	/rpt_family="L2"	17828..17961	
repeatL_region	/rpt_family="L2"	17964..18138	
repeatL_region	/rpt_family="MERL_type"	18139..18209	
repeatL_region	/rpt_family="L2"	18210..18521	
repeatL_region	/rpt_family="Alu"	18522..18865	
repeatL_region	/rpt_family="L2"	19330..19710	
repeatL_region	/rpt_family="L2"	20271..20291	
repeatL_region	/rpt_family="AT-rich"	21212..21327	
repeatL_region	/rpt_family="MIR"	21523..21828	
repeatL_region	/rpt_family="Alu"	22806..22860	
repeatL_region	/rpt_family="Cr-rich"	22994..23105	
repeatL_region	/rpt_family="MIR"	25554..25846	
repeatL_region	/rpt_family="Alu"	25927..26475	
repeatL_region	/rpt_family="L2"	26500..26664	
repeatL_region	/rpt_family="L2"	26754..26897	
repeatL_region	/rpt_family="MERL_type"	26898..27113	
repeatL_region	/rpt_family="L2"	27513..27652	
repeatL_region	/rpt_family="L2"	28002..28057	
repeatL_region	/rpt_family="MERL_type"	28065..28196	
repeatL_region	/rpt_family="Alu"	28481..28541	
repeatL_region	/rpt_family="MERL03"	30093..30138	
repeatL_region	/rpt_family="(TG)n"	33232..33532	
repeatL_region	/rpt_family="Alu"	36502..36816	
repeatL_region	/rpt_family="Alu"	37354..37634	
repeatL_region	/rpt_family="Alu"	41942..41974	
repeatL_region	/rpt_family="L2"	42371..42439	
repeatL_region	/rpt_family="MIR"	42641..42718	
repeatL_region	/rpt_family="(TA)n"	42719..42746	
repeatL_region	/rpt_family="(CA)n"	43843..43966	
repeatL_region	/rpt_family="L2"	47469..47686	
repeatL_region	/rpt_family="MERL_type"	47727..47789	
repeatL_region	/rpt_family="L2"	47840..48153	
Query Match	48.7%;	Score 684.4;	DB 9;
Best Local Similarity	71.1%;	Pred. No. 5.4e-101;	Length 162471;
Matches 1237;	Conservative 0;	Mismatches 76;	Indels 427;
			Gaps

[illegible]

QY 732 GAATTTGCAATCTGAACCAAGTGAAAAACAAATTGCTGAATTGTACTGTATGTAGC 791
|||||
Db 68750 AAATTTGCAATCTGAACCAAGTGAAAAACAAATTGCTGAATTGTACTGTATGTAGC 68691
QY 792 TGCACCTACAAAGATCTTACCCTCTCCACAAGAGGTGAGATGTTAAATGGTCAATCT 851
|||||
Db 68690 TGCACCTACAAAGATCTTACCCTATATCCACAGAGGTGAGATGTTAAATGGTCAATCT 68631
QY 852 GAC--TTTTTTTATTCCTTGACTCAAGACAGTCACTTCATTTTCAGAACCTGTTTA 909
|||
Db 68630 GACTTTTTTTTTTATTTCTCTGTGACTCAAGACAGTCACTTCATTTTCAGAACCTGTTTA 68571
QY 910 AACCTTGTGTGCTGTTTAAATAATGTTGTATACCTGTGTCTTCCTGTATACCA 969
Db 68570 AACCTTGTGTGCTGTTTAAATAATGTTGTATACCTGTGTCTTCCTGTATACCA 68525
QY 970 GACTG--TTTCCCGTGGTGGTTAGATATATTTGTTGATGTTTAAATGGCATGTTT 1028
|||||
Db 68524 GACTGTTTTCGTTGGTGGTTAGATATATTTGTTTGAATGTTTAAATGGTGTGTTT 68465
QY 1029 AGATGTCAGTTTGTAGTCTTCTGAGATGAAGTTCAGCCATTTTGTATCAACAGACAAAG 1088
|||||
Db 68464 AGATGTCAGTATAGTCTTCTGAGATGAAGTTCAGCCATTTTGTATCAACAGACAAAG 68405
QY 1089 CAGTGTCTGTCATTTCCATGCAATAAGTTTGTAGATGTTTAAATGTAAGATCTGATTT 1148
Db 68404 CAGTGTCTGTCATTTCCATGCAATAAGTTTGTAGATGTTTAAATGTAAGATCTGATTT 68345
QY 1149 GCTAGTCTCTCTCTGTAGAGTTAAATGGAAGATTCATCTATCTGTTTAAATGTTCT 1208
|||||
Db 68344 GCTAGTCTCTCTCTGTAGAGTTAAATGGAAGATTCATCTATCTGTTTAAATGTTCT 68285
QY 1209 TCATACCTGTCATTAATTTGGCTGCAGATATTTGTAATTTGTTGCACATCTATGTAAC 1268
Db 68284 TCATACCTGTCATTAATTTGGCTGCAGATATTTGTAATTTGTTGCATCTATGTAAC 68225
QY 1269 AAAACAACCTGAAGATGTTTAAATATATGTAATTTGTAATTTGGAAGTAAAAA 1328
Db 68224 AAAACAACCTGAAGATGTTTAAATATATGTAATTTGTAATTTGGAAGTAAAAA 68165
RESULT 13
MMPT14
LOCUS MMPT14 4678 bp DNA linear ROD 05-APR-1995
DEFINITION Mouse ypc1 gene for ras-related GTP-binding protein, exons 4-6.
ACCESSION X15747
VERSION X15747.1 GI:55463
KEYWORDS GTP-binding protein; ras-related protein; ypc1 gene.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 4678)
AUTHORS Wichmann, H.
TITLE Direct Submission
JOURNAL Submitted (07-JUL-1989) Wichmann H., MPI of Biophysical Chemistry,
Dept of Molecular Genetics, P O Box 2841, D-3400 Goettingen, F R G
REFERENCE 2 (bases 1 to 4678)
AUTHORS Wichmann, H., Disela, C., Haubruck, H. and Galwitz, D.
TITLE Nucleotide sequence of the mouse ypc1 gene encoding a ras-related
GTP-binding protein
JOURNAL Nucleic Acids Res. 17 (16), 6737-6738 (1989)
MEDLINE 89386011
PUBMED 2506528
COMMENT Data kindly reviewed (27-OCT-1989) by Wichmann H.
FEATURES
Location/Qualifiers
1..4678
/organism="Mus musculus"
/strain="c(CH)/c(CH)"
/db_xref="taxon:10090"
precursor_RNA
/note="Primary transcript"

exon 239..334
/note="Exon 4"
/usedin=X15744:gtfb_cds
intron 335..1521
/note="Intron IV"
exon 1522..1653
/note="Exon 5"
/usedin=X15744:gtfb_cds
intron 1654..1802
/note="Intron V"
mRNA 1803..2682
/note="Exon 6"
/note="2000
/note="GTP-binding protein, exon 6 (AA 141-205)"
/usedin=X15744:gtfb_cds
misc_feature 2586..2591
/note="2591
misc_feature 2654..2659
/note="2659
polya_site 2682
/note="polya site"
BASE COUNT 1338 a 913 c 905 g 1522 t
ORIGIN
Query Match 48.6%; Score 682.6; DB 10; Length 4678;
Best Local Similarity 94.8%; Pred. No. 9.5e-101;
Matches 782; Conservative 0; Mismatches 34; Indels 9; Gaps 7;
QY 502 AGGAATTTGCGATTCCTTGGAAATTCGTTTGGAAACAGGCTTAAAGTCAACGA 561
Db 1801 AGGAATTTGCGATTCCTTGGAAATTCGTTTGGAAACAGGCTTAAAGTCAACGA 1860
QY 562 ATGTAGAAGCTCTTTCATGACGATGAGCTGAGATTAAAAAGCAATGGTCCCGAG 621
Db 1861 ATGTAGAAGCTCTTTCATGACGATGAGCTGAGATTAAAAAGCAATGGTCCCGAG 1920
QY 622 CAACAGCTGTGTGTGTGAGAGTCCAAATGTTAAATTCAGACAGCTCCAGTCAAGCT 681
Db 1921 CTACAGCTGTGTGTGTGAGAGTCCAAATGTTAAATTCAGACAGCTCCAGTCAAGCT 1980
QY 682 CAGGTGAGGTTGGCTGCTAAATTTGCCCTCATCTTTTCACAGCAATGAATTTGCA 741
Db 1981 CAGGTGAGGTTGGCTGCTAAATTTGCCCTCATCTTTTCACAGCAATGAATTTGCA 2040
QY 742 TCTGAACCAAGTGAAAAACAAATTCCTGAATTTGCTATGTAAGTGCACATCA 801
Db 2041 TCTGAACCAAGTGAAAAACAAATTCCTGAATTTGCTATGTAAGTGCACATCA 2100
QY 802 CAGATTTCTTACCGTCTCCACAAGGTCCAGATTTGTAATGCTCAATCTGAC--TTTTT 860
Db 2101 CAGATTTCTTACCGTCTTCCAC--AAGGTCCAGATTTGTAATGCTCAATCTTCTTTT 2159
QY 861 TTTATTCCTTGATCCAGACAGCTCACTTCAATTTTCAGAACGTTTAAACCTTTGTG 919
Db 2160 TTTATTCCTTGATCCAGACAGCTCACTTCAATTTTCAGAACGTTTAAACCTTTGTG 2219
QY 920 TGTGTTTAAATAATATGTTGTAATCTTGTGCTTCCCTGATACAGACAGCTTTCC 979
Db 2220 TGTGTTTAAATAATATGTTGTAATCTTGTGCTTCCCTGATACAGACAGCTTTCC 2279
QY 980 CGTGTGTTGTTAGATATATTTGTTTGAATGTTTAAATTTGCAATGTTTGAATGTTG 1039
Db 2280 CGTGTGTTGTTAGATATATTTGTTTGAATGTTTAAATTTGCAATGTTTGAATGTTG 2339
QY 1040 TTAGTCTCTGAAGTGAAGTTCAGCCATTTGTAATCAACAGCAACAGAGTGTGTC 1099
Db 2340 TTAGTCTCTGAAGTGAAGTTCAGCCATTTGTAATCAACAGCAACAGAGTGTGTC 2399
QY 1100 ACTTTCATG--CAATTAAGTTTGTAGATGTTATATGTAAGATGATGATCTGATGTTT 1158
Db 2400 AGTTTCCAGCATTAAGTTTGTAGAGAGC--TATATGTAAGATGATGATGATGATCTT 2458
QY 1159 CTTGTGAAGTTTAAATGAAGATTAACACTATCTGATTAATGATTTCTCATACTCTG 1218

Db 2459 CCGGTAGAGTTAATGAAGACAGACAGACACTATGATTAATTAATTTCTTCACTGCG 2518
 Oy 1219 CATTAAATTTGGTGCAGATATTTGATTTGTCACACTATGTAACAACAACG 1278
 Db 2519 CATTAAATTTGGTGCAGATATTTGATTTGTCACACTATGTAACAACAACG 2574
 Oy 1279 AGATATTTGTAATTAATTAATTTGATTTGTCACACTATGTAACAACAACA 1323
 Db 2575 AGATATTTGTAATTAATTAATTTGATTTGTCACACTATGTAACAACAACA 2619

RESULT 14
 AC117841.C 207945 bp DNA linear HTG 17-JUL-2002
 LOCUS Rattus norvegicus clone CH230-36006, *** SEQUENCING IN PROGRESS
 DEFINITION *** 73 unordered pieces.
 AC117841
 VERSION AC117841.4 GI:21746140
 KEYWORDS HTG: HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 207945)
 Muny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C.,
 Albrooks,S.L., Amaralunga,H.C., Are,J.R., Ayala,M., Banks,T.,
 Barbara,J., Benton,J., Binage,K., Blankenburg,K., Bonini,D.,
 Bouck,J., Bowles,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
 Bulay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denu,A.L., Ding,Y., Dinh,H.H.,
 Donahue,K.R., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,F., Franz,P.,
 Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gattelli,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
 Hernandez,C., Harris,K., Hart,M., Haylak,P., Hayes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,
 Homsl,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
 Jacobson,E., Kelly,S., Khan,U., King,R., Jolivet,S., Joudah,S.,
 Katlovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisged,H.,
 Lozardo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,N., Nickerson,E., Nwokenkwo,S., Ogutu,M., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogutu,M., Nguyen,N.,
 Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Picken,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rivers,M., Rojas,A., Rojubenok,I., Rolfe,M., Ruiz,S., Savary,G.,
 Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Umanu,K., Vasquez,L., Vera,V., Vallalon,D., Vlnson,R., Wang,O.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wlaczek,K., Woodson,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Glibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 207945)
 REFERENCE
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (11-APR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE
 AUTHORS Baylor Plaza, Houston, TX 77030, USA
 TITLE Direct Submission
 JOURNAL Submitted (17-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 14, 2002 this sequence version replaced gi:20258099.
 COMMENT

 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GIBP
 Center clone name: CH230-36006

 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 129522 bases at least Q40
 Consensus quality: 134968 bases at least Q30
 Consensus quality: 138682 bases at least Q20

 NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 73 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1046: contig of 1046 bp in length
 * 1047 1145: gap of unknown length
 * 1147 2183: contig of 1037 bp in length
 * 2184 2283: gap of unknown length
 * 2284 3421: contig of 1138 bp in length
 * 3422 3521: gap of unknown length
 * 3522 4686: contig of 1165 bp in length
 * 4687 4786: gap of unknown length
 * 4787 6158: contig of 1372 bp in length
 * 6159 7327: gap of unknown length
 * 7328 7427: contig of 1069 bp in length
 * 7428 9082: gap of unknown length
 * 9083 9182: contig of 1655 bp in length
 * 9183 10231: gap of unknown length
 * 10232 10331: contig of 1049 bp in length
 * 10332 11663: gap of unknown length
 * 11664 11763: contig of 1332 bp in length
 * 11764 13476: gap of unknown length
 * 13477 13576: contig of 1713 bp in length
 * 13577 15050: gap of unknown length
 * 15051 15150: contig of 1474 bp in length
 * 15151 16522: gap of unknown length
 * 16523 16623: contig of 1372 bp in length
 * 16624 18019: gap of unknown length
 * 18020 18119: contig of 1397 bp in length
 * 18120 19774: gap of unknown length
 * 19775 19874: contig of 1655 bp in length
 * 19875 21203: gap of unknown length
 * 21204 21303: contig of 1329 bp in length
 * 21304 22991: gap of unknown length
 * 22992 23091: contig of 1688 bp in length
 * 23092 24721: gap of unknown length
 * 24722 24821: contig of 1630 bp in length
 * 24822 25833: gap of unknown length
 * 25834 27569: contig of 1012 bp in length
 * 27569 27669: gap of unknown length
 * 27669 29226: contig of 1635 bp in length
 * 29226 29226: contig of 1538 bp in length

*	29227	29326:	gap of unknown length
*	29327	30424:	contig of 1098 bp in length
*	30425	30524:	gap of unknown length
*	30525	32145:	contig of 1621 bp in length
*	32146	32245:	gap of unknown length
*	32246	33498:	contig of 1253 bp in length
*	33499	33598:	gap of unknown length
*	33599	34724:	contig of 1126 bp in length
*	34725	34824:	gap of unknown length
*	34825	36502:	contig of 1678 bp in length
*	36503	36602:	gap of unknown length
*	36603	37971:	contig of 1369 bp in length
*	37972	38071:	gap of unknown length
*	38072	39263:	contig of 1192 bp in length
*	39264	39363:	gap of unknown length
*	39364	41128:	contig of 1765 bp in length
*	41129	41228:	gap of unknown length
*	41229	42650:	contig of 1422 bp in length
*	42651	42730:	gap of unknown length
*	42731	44150:	contig of 1400 bp in length
*	44151	44250:	gap of unknown length
*	44251	45440:	contig of 1190 bp in length
*	45441	45540:	gap of unknown length
*	45541	47375:	contig of 1835 bp in length
*	47376	47475:	gap of unknown length
*	47476	48634:	contig of 1159 bp in length
*	48635	48734:	gap of unknown length
*	48735	50080:	contig of 1346 bp in length
*	50081	50180:	gap of unknown length
*	50181	51471:	contig of 1291 bp in length
*	51472	51571:	gap of unknown length
*	51572	52959:	contig of 1388 bp in length
*	52960	53059:	gap of unknown length
*	53060	54842:	contig of 1783 bp in length
*	54843	54942:	gap of unknown length
*	54943	57384:	contig of 2442 bp in length
*	57385	57484:	gap of unknown length
*	57485	59403:	contig of 1919 bp in length
*	59404	59503:	gap of unknown length
*	59504	60963:	contig of 1460 bp in length
*	60964	61063:	gap of unknown length
*	61064	62330:	contig of 1267 bp in length
*	62331	62430:	gap of unknown length
*	62431	63701:	contig of 1270 bp in length
*	63701	63800:	gap of unknown length
*	63801	65110:	contig of 1310 bp in length
*	65111	65210:	gap of unknown length
*	65211	66895:	contig of 1685 bp in length
*	66896	66995:	gap of unknown length
*	66996	68391:	contig of 1396 bp in length
*	68392	68491:	gap of unknown length
*	68492	70672:	contig of 2181 bp in length
*	70673	70772:	gap of unknown length
*	70773	72373:	contig of 1601 bp in length
*	72374	72473:	gap of unknown length
*	72474	75328:	contig of 2855 bp in length
*	75329	75428:	gap of unknown length
*	75429	77871:	contig of 2443 bp in length
*	77872	77971:	gap of unknown length
*	77972	80743:	contig of 2772 bp in length
*	80744	80843:	gap of unknown length
*	80844	82486:	contig of 1643 bp in length
*	82487	82586:	gap of unknown length
*	82587	84732:	contig of 2146 bp in length
*	84733	84832:	gap of unknown length
*	84833	87219:	contig of 2387 bp in length
*	87220	87319:	gap of unknown length
*	87320	90363:	contig of 3044 bp in length

Query Match 48.3% Score 679.2; DB 2; Length 207945;
 Best Local Similarity 81.9%; Pred. No. 3.8e-100;
 Matches 937; Conservative 0; Mismatches 98; Indels 109; Gaps 9;

Qy	1	AAGCAGAGCTGAGT-GCGGCGCTGCTGATGTGTCTAGGGAGCGAGTAGGAGAGA	59
Db	91854	AAGCAGAGCTGAGTGGGCGGCTGCTGATGTGTATTAAGCGCTGTTGA-GGGCTTGA	91796
Qy	60	CGTTGCTCTCCCGGAACAGCCTATCTCATTCCTTTCTTGAGATTACCCGCGCGAG	119
Db	91795	CGTTGCTCTACTGGAAACAGTATCTTCCTCCCTTCATTAACCGTGGTGCAGAA	91736
Qy	120	AGTCAGGCGCGCGCTGCGAGCAGCAGCGCGCGCGCGCGAGCTGAGTAC	179
Db	91735	GCTTGGGCGCGAGCTGCTGTCAGAAAG---GCGGTGCACGAGAGCAGTGCAGTAC	91679
Qy	180	ATGTCAGCATGAATCCGAATATGATTTATTCAGTTCCTTCGATTGGCGACTCA	239
Db	91678	ATGTCAGCATGAATCCGAATATGATTTATTCAGTTCCTTCGATTGGCGACTCT	91619
Qy	240	GGGTTGGAAGTCTTGCTCTCTTATAGTTTTCAGATGATACATATACGAAGCTAC	299
Db	91618	GGGTTGGAAGTCTTGCTCTCTTATAGTTTTCAGATGATACATATACGAAGCTAC	91559
Qy	300	ATCAGACAA-TTGCTGTGATTTCAAAATAGAAGTATAGTTCAGCGGAGAAACAT	358
Db	91558	ATCAGACATATTTGTGTGTGATTTCAAGATAGGACATAGAGTACAGGAGAAATAT	91499
Qy	359	CAAGCTTCAATA-----	371
Db	91498	CAAGCTTCAATATAGGACACAGCAGCGCAGGAAGATTTGGAACAATACCTCCAGTTA	91439
Qy	372	-----GAGTCCCTTCAA	382
Db	91438	TTACAGAGGAAGCATGTTATCATAGTTGTGTATGATGTGACCGACAGCGAGTCTTCAA	91379
Qy	383	TAATGTAAACAGTGGCTGCTGAGAAATAGATCGTTATGCCAGTAAATGCAACAAT	442
Db	91378	TAAGCTTAACAGTGGCTGCTGAGAAATAGATCGTATGCCAGTAAATGCAACAAT	91319
Qy	443	GTTGTAGGGAACAATGTGATCTGACCACAAGAAAGTATGACTACACACAGGAA	502
Db	91318	GTTGTAGGGAACAACAATGTGATCTGACAAAC---AAAGTATGATCTACACAAAGCAA	91263
Qy	503	GGAATTCGATTCCTTCGTAATTCGTTTGGAAACCGATGCTTAAGATGACAGAA	562
Db	91262	GGAATTCGATTCCTTCGTAATTCGTTTGGAAACCGATGCTTAAGATGACAGAA	91203
Qy	563	TGTGAACAGCTTTTCATGAGATGCGAGTATGATTTAAAGGAATGGTCCCGAGC	622
Db	91202	TGTGAACAGCTTTTCATGAGATGCGAGTATGATTTAAAGGAATGGTCCCGAGC	91143
Qy	623	AACAGCTGTGTGTGAGAAATGCAATGTTAAATTCAGACACCTCCAGTCAAGAGTC	682
Db	91142	AACAGCTGTGTGTGAGAAATGCAATGTTAAATTCAGACACCTCCAGTCAAGAGTC	91083
Qy	683	AGTGTGAGTGTGTGTGAAATTTTCCCATCTCTTTCACAGCAATGAAATTTGCAAT	742
Db	91082	AGTGTGAGTGTGTGTGAAATTTTCCCATCTCTTTCACAGCAATGAAATTTGCAAT	91023
Qy	743	CTGAACCCAGTGAAGAAACAAAT-TGCTGAATGTACTGTATGAGTGCAGTACAA	801
Db	91022	CTGAACCCAGTGAAGAAACAAATCTGCAATGTACTGTATGAGTGCAGTACAA	90963
Qy	802	CAGATTTCTTACCGTCTCCACAAGGTGAGATGTAATGTCATATCTGAC-TTTTTT	860
Db	90962	CAGATTTCTTACCGTCTTTCAC-AAGGTGAGATGTAATGTAATGTAATCTGACTTTT	90904
Qy	861	TTTATTCCTTGCAGTCAAGAGTAACTATTTTGAACCTGTTTAACTTTGTGT	920
Db	90903	TTTATTCCTTGCAGTCAAGAGTAACTATTTTGAACCTGTTTAACTTTGTGT	90844
Qy	921	GCTGTTTATTAATTAATGTGTATATCTGTTGCTTCCCTGATACAGACTGTTTCCC	980
Db	90843	GCTGTTTATTAATTAATGTGTATATCTGTTGCTTCCCTGATACAGACTGTTTCCC	90784
Qy	981	GTGTTGTGTTAAGATATATTTTGTGTGATGTTATATGTCAGATTTAGATGTCAGTT	1040

|||||
 Db 90783 GTCGTTTGTGAAATATTTTCTTTCATCTTATATTCGACGTTTACATCTTGCGTT 90724
 QY 1041 TAGT 1044
 Db 90723 TAGT 90720
 RESULT 15
 AC099354 174028 bp DNA linear HTG 11-JUL-2002
 LOCUS Rattus norvegicus clone CH230-901, *** SEQUENCING IN PROGRESS ***
 DEFINITION 67 unordered pieces.
 ACCESSION AC099354
 VERSION AC099354.3 GI:21722818
 KEYWORDS HTG, HTGS, PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Rodentia; Muridae; Murinae;
 Rattus
 1 bases 1 to 174028)
 1 Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
 Alshrocks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T.,
 Barbarella, J., Benton, J., Blomberg, K., Blankenburg, K., Bonnin, D.,
 Bouch, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Burch, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, F., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.B., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthett, K.J., Draper, H., Dugan-Rocha, S., Dunblin, K.J.,
 Ebnharth, C., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
 Homsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Lounsgood, H.,
 Lopez, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, P.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, N., Okunolu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, S., Saeed, G.,
 Rivers, M., Rojas, A., Rojuben, I., Rolfe, M., Ruiz, S., Saverly, G.,
 Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, J.,
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tameris, A., Tameris, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, V., Telford, B., Thomas, R., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, K., Wang, C.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Walling, K.,
 Williams, G., Williamson, A., Wleczek, R., Woodson, C., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 TITLE Unpublished
 JOURNAL Direct Submission
 REFERENCE 2 (bases 1 to 174028)
 AUTHORS Worley, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 174028)
 AUTHORS Worley, K.C.
 TITLE Direct Submission

JOURNAL COMMENT

Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 10, 2002 this sequence version replaced gi:17942431.

 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: CH230-901
 Center clone name: CH230-901

 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap version 0.990329
 Consensus quality: 121217 bases at least Q40
 Consensus quality: 127132 bases at least Q30
 Consensus quality: 131924 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 67 contigs in the true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N's, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1
 1005 1004: contig of 1004 bp in length
 1105 1104: gap of unknown length
 2302 2301: contig of 1197 bp in length
 2402 2401: gap of unknown length
 3457 3456: contig of 1053 bp in length
 3557 3556: gap of unknown length
 4595 4594: contig of 1038 bp in length
 4695 4694: gap of unknown length
 5599 5598: contig of 1205 bp in length
 6000 5999: gap of unknown length
 7392 7391: contig of 1392 bp in length
 7492 7491: gap of unknown length
 7992 7991: contig of 1072 bp in length
 8563 8562: gap of unknown length
 8664 8663: contig of 1013 bp in length
 9677 9676: gap of unknown length
 9777 9776: gap of unknown length
 11544 11543: contig of 1767 bp in length
 11644 11643: gap of unknown length
 12809 12808: contig of 1165 bp in length
 12909 12908: gap of unknown length
 14321 14320: contig of 1412 bp in length
 14421 14420: gap of unknown length
 15913 15912: contig of 1493 bp in length
 16014 16013: gap of unknown length
 17145 17144: contig of 1132 bp in length
 17245 17244: gap of unknown length
 18550 18549: gap of 1305 bp in length
 18651 18650: gap of unknown length
 20071 20070: contig of 1420 bp in length
 20171 20170: gap of unknown length
 21207 21206: contig of 1036 bp in length
 21307 21306: gap of unknown length
 22659 22658: contig of 1352 bp in length
 22759 22758: gap of unknown length
 24736 24735: contig of 1977 bp in length
 24836 24835: gap of unknown length
 26166 26165: contig of 1330 bp in length
 26266 26265: gap of unknown length
 28440 28439: contig of 2173 bp in length
 28441 28440: gap of unknown length
 30486 30485: contig of 1943 bp in length
 30586 30585: gap of unknown length
 31765: contig of 1180 bp in length

31766	31865	gap of unknown length
31866	32886	contig of 1021 bp in length
32887	32986	gap of unknown length
32987	34733	contig of 1747 bp in length
34734	34833	gap of unknown length
34834	35561	contig of 1728 bp in length
35562	36661	gap of unknown length
36662	38855	contig of 2194 bp in length
38856	38955	gap of unknown length
38956	40675	contig of 1720 bp in length
40676	40775	gap of unknown length
40776	43313	contig of 2538 bp in length
43314	43413	gap of unknown length
43414	45970	contig of 2557 bp in length
45971	46070	gap of unknown length
46071	48910	contig of 2840 bp in length
48911	49010	gap of unknown length
49011	51895	contig of 2885 bp in length
51896	51995	gap of unknown length
51996	54123	contig of 2128 bp in length
54124	54223	gap of unknown length
54224	55763	contig of 1540 bp in length
55764	55863	gap of unknown length
55864	57512	contig of 1649 bp in length
57513	57612	gap of unknown length
57613	59522	contig of 1910 bp in length
59523	59622	gap of unknown length
59623	60937	contig of 1315 bp in length
60938	61037	gap of unknown length
61038	63206	contig of 2169 bp in length
63207	63306	gap of unknown length
63307	64956	contig of 1650 bp in length
64957	65056	gap of unknown length
65057	67731	contig of 2675 bp in length
67732	67831	gap of unknown length
67832	70617	contig of 2786 bp in length
70618	70717	gap of unknown length
70718	73447	contig of 2730 bp in length
73448	73547	gap of unknown length
73548	77162	contig of 3615 bp in length
77163	77262	gap of unknown length
77263	80487	contig of 3225 bp in length
80488	80587	gap of unknown length
80588	82852	contig of 2265 bp in length
82853	82952	gap of unknown length
82953	86851	contig of 3899 bp in length
86852	86951	gap of unknown length
86952	91799	contig of 4848 bp in length
91800	91899	gap of unknown length
91900	95211	contig of 3312 bp in length
95212	95311	gap of unknown length
95313	98112	contig of 2801 bp in length
98113	98212	gap of unknown length
98213	100745	contig of 2533 bp in length
100746	100845	gap of unknown length
100846	104055	contig of 3210 bp in length
104056	104155	gap of unknown length
104156	107669	contig of 3514 bp in length
107670	107768	gap of unknown length
110770	111040	contig of 3271 bp in length
111041	111140	gap of unknown length
111437	114378	contig of 3238 bp in length
114379	114478	gap of unknown length
114479	117845	contig of 3367 bp in length

Query Match 40.88; Score 573.8; DB 2; Length 174028;
 Best Local Similarity 79.98; Pred. No. 3,7e-83;
 Matches 822; Conservative 0; Mismatches 97; Indels 110; Gaps 8;

OY 1 AAGGATACGAGT-GGGGGGCTGCTGATGTTGTTAGGAGCAGAGTGGGAGA 59
 Db 100711 AAGGATACGAGTGGGGGCTGCTGATGTTGTTAGTGTGAGTGGGTGA 100653

OY 60	CGTTTCTCCCGGACAGCCTATCTCATTCCTTTCTTTGATTAACCGTGGCGGAG 119
Db 100652	CGTTTCTCTACTGGAACAGCTTATGCTCATTCCTTCCCTTCATTAACATGTTGACGA 100593
OY 120	AGTACGGGCGGCGCTGGGACAGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 179
Db 100592	GCTTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 100536
OY 180	ATGTCACATGAATCCGGAATGATGATTTATTTATTCAGTACTTCTGATGGGACATCA 239
Db 100555	ATGTCACATGAATCCGGAATGATGATTTATTTATTTATTTATTTATTTATTTATTTATTT 100479
OY 240	GGGGTTGGAAGCTTGGCTTCTTCTTATGTTTGCAGATGATATACAGAAAGCTAC 299
Db 100478	GGGGTTGGAAGCTTGGCTTCTTCTTATGTTTGCAGATGATATACAGAAAGCTAC 100419
OY 300	ATCAGCAATTTGGTGTGATTTTCAATTAAGACTATAGATTAAGAGGAGGAGGAGGAGGAGG 359
Db 100418	ATCAGCAATTTGGTGTGATTTTCAATTAAGACTATAGATTAAGAGGAGGAGGAGGAGGAGG 100359
OY 360	AAGCTTCAATA----- 371
Db 100358	AAGCTTCAATA----- 371
OY 372	-----GAGTCTTCAAT 383
Db 100298	TACAGAGAACCATGCGCATGATGTTGTATGATGATGACGACAGACAGGACCTTCAAT 100239
OY 384	AATGTTAAACAGTGGCTCAGAGAAATGATGCTTATGCTATGAAATGTTAAACAAATG 443
Db 100238	AATGTTAAACAGTGGCTCAGAGAAATGATGCTTATGCTATGAAATGTTAAACAAATG 100179
OY 444	TTGGTAGGGAACAATGATGATCTGACCAACAAGAGTAGACTACACAACAGCAAG 503
Db 100178	TTGGTAGGGAACAATGATGATCTGACCAACAAGAGTAGACTACACAACAGCAAG 100123
OY 504	GAATTTGCTGATTCCTTGAATTCGTTTGGAAACAGTGTAGAAATGCAACGAT 563
Db 100122	GAATTTGCTGATTCCTTGAATTCGTTTGGAAACAGTGTAGAAATGCAACGAT 100063
OY 564	GTAAGACAGTCTTTCATGACATGCGACCTGAGATTAATAAGCAATGGTCCGGAGCA 623
Db 100062	GTAAGACAGTCTTTCATGACATGCGACCTGAGATTAATAAGCAATGGTCCGGAGCA 100003
OY 624	ACAGTGTGCTGCTGAGAACTCCAAATGTTAAATTCAGACACCTCCAGTAAAGCA 683
Db 100002	ACAGTGTGCTGCTGAGAACTCCAAATGTTAAATTCAGACACCTCCAGTAAAGCA 99943
OY 684	GGTGAGGCTGCTCTAAATTTGCTCCATCCTTTCTCAGACCAATGAATTTGCAATC 743
Db 99942	GGTGAGGCTGCTCTCTAAATTTGCTCCATCCTTTCTCAGACCAATGAATTTGCAATC 99883
OY 744	TGAACCCAGTGAAGAAACAAATTTGCTGATTTGATGATGATGATGATGATGATGATGATG 803
Db 99882	TGAACCCAGTGAAGAAACAAATTTGCTGATTTGATGATGATGATGATGATGATGATGATG 99823
OY 804	GATTTCTACGCTGCTCAGAAAGTCCAGATGTAATGTAATGTAATGTAATGTAATGTAATG 862
Db 99822	GATTTCTACGCTGCTCAGAAAGTCCAGATGTAATGTAATGTAATGTAATGTAATGTAATG 99764
OY 863	TATTCCTTGAAGTCAAGACAGTCAATTTTCAAGAACTGTTTAAACCTTTGTGTGC 922
Db 99763	TATTCCTTGAAGTCAAGACAGTCAATTTTCAAGAACTGTTTAAACCTTTGTGTGC 99704
OY 923	TGTTTATA 931
Db 99703	TGTTTATA 99695

Search completed: November 17, 2002, 12:50:42
 Job time : 4075 secs

DR WPI; 2000-594639/56.
 DR P-PSDB; AAB34816.
 XX Fifty nucleic acid molecules encoding human secreted proteins, useful
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases -
 PS
 XX Claim 1; Page 371; 425p; English.
 CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the
 CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
 CC AAB34852 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities include:
 CC cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antifungal; antiparasitic; and
 CC vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for
 CC preventing, treating or ameliorating a medical condition in e.g. humans,
 CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities. The polynucleotides are
 CC useful for chromosome identification. They are also useful as probes for
 CC diagnosing a disorder related to the female reproductive system,
 CC particularly breast and/or ovary cancer. They are also useful in the gene
 CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
 CC agonists and antagonists from the present invention are useful in the
 CC diagnosis, treatment and prevention of cancer, immune disorders,
 CC cardiovascular disorders, wound healing, neurological diseases and
 CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 2528 BP; 772 A; 428 C; 545 G; 781 T; 2 other;

Query Match 85.48; Score 1200.2; DB 21; Length 2528;
 Best Local Similarity 92.96; Pred. No. 4.3e-155;
 Matches 1319; Conservative 0; Mismatches 4; Indels 97; Gaps 2;

QY 1 AAGCGATAGCTGAGT-GGGCGGCTGCTGATGTTGTTCTAGGGGAGGAGGAGGA 59
 Db 28 AAGCGATAGCTGAGTGGCGGCGGCTGCTGATGTTGTTCTAGGGGAGGAGGAGGA 87
 QY 60 CGTTTGCTCTCCGGAACAGCCTATCTCATCTCTTCTTGATTACCGGTGGCGGAG 119
 Db 88 CGTTTGCTCTCCGGAACAGCCTATCTCATCTCTTCTTGATTACCGGTGGCGGAG 147
 QY 120 AGTAGGGCGGCGGCTGCGGAGCAAGGGCGGCTGGCGGCGGCGAGCTGCAGTGC 179
 Db 148 AGTAGGGCGGCGGCTGCGGAGCAAGGGCGGCTGGCGGCGGCGAGCTGCAGTGC 207
 QY 180 ATGCCAGCATGATCCCGAATATGATTATTTCAAGTACTCTGATGGGAGCTCA 239
 Db 208 ATGCCAGCATGATCCCGAATATGATTATTTCAAGTACTCTGATGGGAGCTCA 267
 QY 240 GGGGTGGAAAGTCTTCCCTCTCTTCTAGTTTCAGATGATACATATACGAAGCTAC 299
 Db 268 GGGGTGGAAAGTCTTCCCTCTCTTCTAGTTTCAGATGATACATATACGAAGCTAC 327
 QY 300 ATCCAGCAATTTGGTGTGGATTTCAAAATAGAACTATAGAGTTCAGGGAACAATC 359
 Db 328 ATCCAGCAATTTGGTGTGGATTTCAAAATAGAACTATAGAGTTCAGGGAACAATC 387
 QY 360 AAGCTCAATA----- 371
 Db 388 AAGCTCAATAATATGGACACAGCAGCCAGAAAGATTTGCAACAATCCTCCAGTTAT 447
 QY 372 -----GAGTCCCTCAAT 383
 Db 448 TACAGAGAGCCATGGATCATAGTTGTATGATGAGACAGATTCAGGAATCCTTCAAT 507
 QY 384 AATGTTAAACAGTGGCTGACGAGAAATAGATGTTATGCCAGTGAATAATGTCAACAAATTG 443
 XX

Db 508 AATGTTAAACAGTGGCTGACGAGAAATAGATGTTATGCCAGTGAATAATGTCAACAAATTG 567
 QY 444 TTGTTAGGAACAATGTGATCTGACCAACAAGAAATAGTATGACTACACAACGCGAAG 503
 Db 568 TTGTTAGGAACAATGTGATCTGACCAACAAGAAATAGTATGACTACACAACGCGAAG 627
 QY 504 GAATTTGCTGATTCCTTGGAATTCGTTTTGGAAACAGTGTCTAAGAAATGCACGAAT 563
 Db 628 GAATTTGCTGATTCCTTGGAATTCGTTTTGGAAACAGTGTCTAAGAAATGCACGAAT 687
 QY 564 GTAAACAGTCTTCATGACATGACATGACGCTGAGTTTAAAGCAATGGGTCGCGGAGCA 623
 Db 688 GTAAACAGTCTTCATGACATGACATGACGCTGAGTTTAAAGCAATGGGTCGCGGAGCA 747
 QY 624 ACAGCTGGTGTGTGAGAAAGTCCAAATGTTAAATTCAGACACTCCAGTCAACAGTCA 683
 Db 748 ACAGCTGGTGTGTGAGAAAGTCCAAATGTTAAATTCAGACACTCCAGTCAACAGTCA 807
 QY 684 GGTGAGGTTGCTGCTAAATTTGCTCCATCCTTTCTCAGACAAATGAAATTTGCAATC 743
 Db 808 GGTGAGGTTGCTGCTAAATTTGCTCCATCCTTTCTCAGACAAATGAAATTTGCAATC 867
 QY 744 TGAACCAAGTGAAGAAACAAATTTGCTGATTTGATGTACTGTATGACTGACTACACA 803
 Db 868 TGAACCAAGTGAAGAAACAAATTTGCTGATTTGATGTACTGTATGACTGACTACACA 927
 QY 804 GATCTTACGCTCCCAAAAGTCCAGAGATTTGAAATGGCAATGACTGACTTTTCTTTT 863
 Db 928 GATCTTACGCTCCCAAAAGTCCAGAGATTTGAAATGGCAATGACTGACTTTTCTTTT 987
 QY 864 ATTCCTTGACTCAAGACAGCTAACTTCATTTTCAGAACTGTGTTTAAACCTTGTGTCT 923
 Db 988 ATTCCTTGACTCAAGACAGCTAACTTCATTTTCAGAACTGTGTTTAAACCTTGTGTCT 1047
 QY 924 GGTATTAATAATATGTGTGTATTCCTTTGCTTCTGCTGATACCAACGTTCCCGTG 983
 Db 1048 GGTATTAATAATATGTGTGTATTCCTTTGCTTCTGCTGATACCAACGTTCCCGTG 1107
 QY 984 GTTGTGTAATATATTTGTTTGTATGTTATATTTGGCATGTTTATGATGTCAGGTTTAA 1043
 Db 1108 GTTGTGTAATATATTTGTTTGTATGTTATATTTGGCATGTTTATGATGTCAGGTTTAA 1167
 QY 1044 TCTTCTGAAGTGAAGTTCAGCCATTTTGTATCAACAGCAACAGAGTGTCTGTCACTT 1103
 Db 1168 TCTTCTGAAGTGAAGTTCAGCCATTTTGTATCAACAGCAACAGAGTGTCTGTCACTT 1227
 QY 1104 TCCATGATATAAGTTTGTAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1163
 Db 1228 TCCATGATATAAGTTTGTAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1287
 QY 1164 TAGAGTATATAATGGAAGATTTACACTATCTGATTAATAGTTTCTTCTATCTGCTATAT 1223
 Db 1288 TAGAGTATATAATGGAAGATTTACACTATCTGATTAATAGTTTCTTCTATCTGCTATAT 1347
 QY 1224 AATTTGCTGCGAATATTTGTAATTTGTTGACACACTATGTAACAAACACTGAAGAT 1283
 Db 1348 AATTTGCTGCGAATATTTGTAATTTGTTGACACACTATGTAACAAACACTGAAGAT 1407
 QY 1284 AATGTTAATAATATTTGATCTTATTTGGAAGTAAATAAAAA 1323
 Db 1408 AATGTTAATAATATTTGATCTTATTTGGAAGTAAATAAAAA 1447
 XX
 RESULT 2
 AAS83860/C
 ID AAS83860 standard; cDNA; 1193 BP.
 AC AAS83860;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #19664.
 XX

KH Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB: ABG19673.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1: SEQ ID No 19664; 103pp: English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on human
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 1193 BP: 372 A; 279 C; 213 G; 329 T; 0 other;
SO
Query Match 73.2%; Score 1029; DB 23; Length 1193;
Best Local Similarity 89.9%; Pred. No. 8.7e-132;
Matches 1182; Conservative 0; Mismatches 0; Indels 133; Gaps 2;

Db 943 |||||
GGGCTTGAAAGCTCTTGCTCTTCTTAGGTTGGAGTGTATACATATACGAAAGCTAC 884
QY 300 ATCAGCACAATTTGGTGTGATTTCAAAATTAAGAACTATAGAGTTAGACGGGAAACATC 359
Db 883 ATCAGCACAATTTGGTGTGATTTCAAAATTAAGAACTATAGAGTTAGACGGGAAACATC 824
QY 360 AAGCTTCAAAATAGAGTCTCTTCAATTAATGTAAACAGTGGCTGCAGAAATAGATCTTAT 419
Db 823 AAGCTTCAAAAT- 812
QY 420 GCCAGTGAATAATGTCAACAAATTTGTTGAGGAACAATGTATCTGACCAAGAA 479
Db 811 ----- 812
QY 480 GTAGTACTACACACAGGAGAAATTTGCTGATTCCTTGGAATTCGTTTTGGAA 539
Db 811 -----GAAATTTGCTGATTCCTTGGAATTCGTTTTGGAA 776
QY 540 ACCAGTCTAAGAAATGCAAGAGAAATTTGCTGATTCCTTGGAATTCGTTTTGGAA 599
Db 775 ACCAGTCTAAGAAATGCAAGAGAAATTTGCTGATTCCTTGGAATTCGTTTTGGAA 716
QY 600 AAAAAGCAATGGGTCCCGGAGCAACAGCTGGTGTCTGAGAAATTCATTTAAAT 659
Db 715 AAAAAGCAATGGGTCCCGGAGCAACAGCTGGTGTCTGAGAAATTCATTTAAAT 656
QY 660 CAGAGCACTCCAGTCAACAGAGTCAAGTGGAGGTTGCTGAATTTGCTCCATCCTT 719
Db 655 CAGAGCACTCCAGTCAACAGAGTCAAGTGGAGGTTGCTGAATTTGCTCCATCCTT 596
QY 720 TCTCAGCAATGAATTTGCAATCTGAAACCAAGTAAACAAATTTGCTCAATCTT 779
Db 595 TCTCAGCAATGAATTTGCAATCTGAAACCAAGTAAACAAATTTGCTCAATCTT 536
QY 780 ACTGTATAGTGTGACATACACAGATTTCTACGCTCTCCACAAGGTCAGAGATTGA 839
Db 535 ACTGTATAGTGTGACATACACAGATTTCTACGCTCTCCACAAGGTCAGAGATTGA 476
QY 840 ATGTGTCAATCTGACTTTTATTTTATTTCCCTGACATCAAGAGTCAATCTTTCAG 899
Db 475 ATGTGTCAATCTGACTTTTATTTTATTTCCCTGACATCAAGAGTCAATCTTTCAG 416
QY 900 AACTGTTTAAACCTTTGTGCTGGTTTAAATTAATGTGTATCTGTTGCTT 959
Db 415 AACTGTTTAAACCTTTGTGCTGGTTTAAATTAATGTGTATCTGTTGCTT 356
QY 960 CCGATATCAGAGCTTTCCCGGAGTGGTGTGAATAATTTGTTGATTTATAT 1019
Db 355 CCGATATCAGAGCTTTCCCGGAGTGGTGTGAATAATTTGTTGATTTATAT 296
QY 1020 GCGATGTTAGATGCTGAGTGTAGCTCTGTAAGTGAAGTCAAGCATTTGATTA 1079
Db 295 GCGATGTTAGATGCTGAGTGTAGCTCTGTAAGTGAAGTCAAGCATTTGATTA 236
QY 1080 CAGCAACAGAGTGTCTGCTCACTTCCATGCAATAAGTTAGAGATGTTATGTA 1139
Db 235 CAGCAACAGAGTGTCTGCTCACTTCCATGCAATAAGTTAGAGATGTTATGTA 176
QY 1140 ATCTGATTGTAGTCTTCTCTGTAGAGTTAATAATGAAAGATTACATCTGATTA 1199
Db 175 ATCTGATTGTAGTCTTCTCTGTAGAGTTAATAATGAAAGATTACATCTGATTA 116
QY 1200 ATATTTCTTCAATCTGCAATAATAATTTGCTGCAATAATTTGTTGCA 1259
Db 115 ATATTTCTTCAATCTGCAATAATAATTTGCTGCAATAATTTGTTGCA 56
QY 1260 CTATGTAAACAAACAACTGAAGATATGTTAATAATTTGCTATTTGGAAGT 1314
Db 55 CTATGTAAACAAACAACTGAAGATATGTTAATAATTTGCTATTTGGAAGT 1

RESULT 3

ABK44823/C
ID ABK44823 standard; cDNA: 506 BP.
XX
AC ABK44823;
XX
DE 05-JUN-2002 (first entry)
XX
DE cDNA encoding colon tumour protein, SEQ ID No 374.
XX
KW Human: colon tumour; vaccine; colon cancer; immunogenic;
KW Immunotherapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200212328-A2.
XX
PD 14-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US24218.
XX
PR 03-AUG-2000; 2000US-223283P.
PR 28-MAR-2001; 2001US-279763P.
PR 29-JUN-2001; 2001US-302051P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Kling GE, Meagher MJ, Xu J, Secretist H;
XX
DR WPI; 2002-241739/29.
XX
PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
PT for diagnosing, preventing, and treating colon cancer, and as markers
PT for the progression of cancer -
XX
PS Claim 1; SEQ ID No 374; 147pp: English.
XX
CC The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
CC compositions, such as vaccines, for the diagnosis, prevention, and
CC treatment of colon cancer. Polynucleotide sequences may be used as
CC hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. The compositions are useful for stimulating an
CC immune response against cancer, particularly for the immunotherapy of
CC colon cancer, and as markers for the progression of cancer.
CC ABK44450-ABK46237 represent coding sequences of human colon tumour
CC proteins of the invention.
CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
CC for this patent did not form part of the printed specification but was
CC supplied by the European Patent Office.
XX
SQ Sequence 506 BP; 194 A; 91 C; 82 G; 139 T; 0 other;
Query Match 36.0%; Score 506; DB 24; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.2e-60;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 770 CCTGAATGTACTGTATGACCTGACACAAAGATCTTACCGTCCCAAGGTC 829
DB 506 CCTGAATGTACTGTATGACCTGACACAAAGATCTTACCGTCCCAAGGTC 447
QY 830 GAGATTTGTAATGTCATGACTGATTTTATTCCTTGAAGTCAAGACGTA 889
DB 446 GAGATTTGTAATGTCATGACTGATTTTATTCCTTGAAGTCAAGACGTA 387
QY 890 TCATTTTCAGAACTGTTTAAACCTTTGTGCTGCTTTATTAATATGTTGTAATCC 949
DB 386 TCATTTTCAGAACTGTTTAAACCTTTGTGCTGCTTTATTAATATGTTGTAATCC 327
QY 950 TTGTTGCTTCTGATACCAAGCTGTTCCGCTGGTGGTGAAGATATTTGTTGA 1009
DB 326 TTGTTGCTTCTGATACCAAGCTGTTCCGCTGGTGGTGAAGATATTTGTTGA 267

QY 1010 TGTATATATGCGATGTTTACAGTGTGAGTTTACTCTTCTGAAGATGAAGTTCAGCCATT 1069
DB 266 TGTATATATGCGATGTTTACAGTGTGAGTTTACTCTTCTGAAGATGAAGTTCAGCCATT 207
QY 1070 TTGATCAAAACAGCAAGCAGTGTCTGCTCACTTCCATGATTAAGTTTGTAGATCT 1129
DB 206 TTGATCAAAACAGCAAGCAGTGTCTGCTCACTTCCATGATTAAGTTTGTAGATCT 147
QY 1130 TATATGTAAGATGATGTTGCTAGTCTTCCCTGTAGAGTTATTAATGGAAGATTACAC 1189
DB 146 TATATGTAAGATGATGTTGCTAGTCTTCCCTGTAGAGTTATTAATGGAAGATTACAC 87
QY 1190 TATCTGATTAATGATTTCTTCATCTGCAATATTAATTTGTCGTCAGAAATATTGTAAT 1249
DB 86 TATCTGATTAATGATTTCTTCATCTGCAATATTAATTTGTCGTCAGAAATATTGTAAT 27
QY 1250 TTGTTGCACACTATGTATCAAAACAA 1275
DB 26 TTGTTGCACACTATGTATCAAAACAA 1
RESULT 4
ABK83978
ID ABK83978 standard; cDNA: 723 BP.
XX
AC ABK83978;
XX
DE 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #549.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; AIDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 549; 114pp: English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) Gs by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the

Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.2%; Score 467; DB 22; Length 487;
Best Local Similarity 100.0%; Pred. No. 2,4e-55;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 AGGAATTGCTGATTCCTTGGAAATTCCTTTTGGAAACCAAGCTGCTAAGATCAACGA 561
DB 467 AGGAATTGCTGATTCCTTGGAAATTCCTTTTGGAAACCAAGCTGCTAAGATCAACGA 408
QY 562 ATGAGAACAGCTCTTTCATGACGATGACGATGATTAATAACGATGGGTCGCGAG 621
DB 407 ATGAGAACAGCTCTTTCATGACGATGACGATGATTAATAACGATGGGTCGCGAG 348
QY 622 CAACAGCTGCTGCTGAGAGATCAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 681
DB 347 CAACAGCTGCTGCTGAGAGATCAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 288
QY 682 CAGGTGAGGCTGCTGCTAATAATTCCTTCATCCTTTTCTCAGACGATTAATTTGCA 741
DB 287 CAGGTGAGGCTGCTGCTAATAATTCCTTCATCCTTTTCTCAGACGATTAATTTGCA 228
QY 742 TCTGAACCAAGTGAATAATTCCTGAAATGCTATGATGATGATGATGATGATGATGAT 801
DB 227 TCTGAACCAAGTGAATAATTCCTGAAATGCTATGATGATGATGATGATGATGATGAT 168
QY 802 CAGATTTCTACCGTCTCCACAAAGTCAAGATTTGAATGCTAATGCTACTTTT 861
DB 167 CAGATTTCTACCGTCTCCACAAAGTCAAGATTTGAATGCTAATGCTACTTTT 108
QY 862 TTATTCCTTGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 921
DB 107 TTATTCCTTGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 48
QY 922 CTGGTTTATAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 968
DB 47 CTGGTTTATAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1

RESULT 6
ABA53512/c
ID ABA53512 standard; DNA; 487 BP.

AC ABA53512;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #1817.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN MO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001MO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 1; SEQ ID NO 1817; 639pp + sequence listing; English.
PS

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX

Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.2%; Score 467; DB 22; Length 487;
Best Local Similarity 100.0%; Pred. No. 2,4e-55;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 AGGAATTGCTGATTCCTTGGAAATTCCTTTTGGAAACCAAGCTGCTAAGATCAACGA 561
DB 467 AGGAATTGCTGATTCCTTGGAAATTCCTTTTGGAAACCAAGCTGCTAAGATCAACGA 408
QY 562 ATGAGAACAGCTCTTTCATGACGATGACGATGATTAATAACGATGGGTCGCGAG 621
DB 407 ATGAGAACAGCTCTTTCATGACGATGACGATGATTAATAACGATGGGTCGCGAG 348
QY 622 CAACAGCTGCTGCTGAGAGATCAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 681
DB 347 CAACAGCTGCTGCTGAGAGATCAATGTTAAATTCAGAGCACTCCAGTCAAGCAGT 288
QY 682 CAGGTGAGGCTGCTGCTAATAATTCCTTCATCCTTTTCTCAGACGATTAATTTGCA 741
DB 287 CAGGTGAGGCTGCTGCTAATAATTCCTTCATCCTTTTCTCAGACGATTAATTTGCA 228
QY 742 TCTGAACCAAGTGAATAATTCCTGAAATGCTATGATGATGATGATGATGATGATGAT 801
DB 227 TCTGAACCAAGTGAATAATTCCTGAAATGCTATGATGATGATGATGATGATGATGAT 168
QY 802 CAGATTTCTACCGTCTCCACAAAGTCAAGATTTGAATGCTAATGCTACTTTT 861
DB 167 CAGATTTCTACCGTCTCCACAAAGTCAAGATTTGAATGCTAATGCTACTTTT 108
QY 862 TTATTCCTTGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 921
DB 107 TTATTCCTTGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 48
QY 922 CTGGTTTATAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 968
DB 47 CTGGTTTATAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1

RESULT 7
ABA23276/c
ID ABA23276 standard; DNA; 487 BP.

AC ABA23276;

DT 23-JAN-2002 (first entry)

DE Probe #1742 for gene expression analysis in human heart cell sample.

KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.

OS Homo sapiens.

PN MO200157274-A2.

PD 09-AUG-2001.

Db 227 TCTAACCACGAGTAAACAAATGCGTGAATGTGATGTAAGTGCACACTACAA 168
 Qy 802 CAGATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 861
 Db 167 CAGATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 108
 Qy 862 TTATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 921
 Db 107 TTATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 48
 Qy 922 CTGCTTTATTAATAATGATGTGTATCTTGTCTTCTTCCGATACC 968
 Db 47 CTGCTTTATTAATAATGATGTGTATCTTGTCTTCCGATACC 1

RESULT 9
 AAK27239/C
 ID AAK27239 standard; DNA: 487 BP.

AC AAK27239;
 DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 1796.
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for
 analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 1796; 658bp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.

XX Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.28; Score 467; DB 22; Length 487;

Best Local Similarity 100.0%; Pred. No. 2,4e-55;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 502 AGGAATTTGGGATTCCTTGGATTCGTTTGGAAACGAGCTAAGATGCAACGA 561
 Db 467 AGGAATTTGGGATTCCTTGGATTCGTTTGGAAACGAGCTAAGATGCAACGA 408

Oy 562 ATGTAGAACAGCTCTTTCATGACGATGGCAGCTGAGATTTAAAGCGAATGGGTCCGAG 621
 Db 407 ATGTAGAACAGCTCTTTCATGACGATGGCAGCTGAGATTTAAAGCGAATGGGTCCGAG 348
 Qy 622 CAACAGCGGTGGGTGCTGAGAGTCCAAATGTTAAATTCAGAGCTCCAGTCAAGCAGT 681
 Db 347 CAACAGCGGTGGGTGCTGAGAGTCCAAATGTTAAATTCAGAGCTCCAGTCAAGCAGT 288
 Qy 682 CAGGTGAGAGTGGTGTCTAAATTTGGCTCCATCCCTTTCTCAGAGCAATGAATTTGCA 741
 Db 287 CAGGTGAGAGTGGTGTCTAAATTTGGCTCCATCCCTTTCTCAGAGCAATGAATTTGCA 228
 Qy 742 TCTAACCACGAGTAAACAAATGCGTGAATGTGATGTAAGTGCACACTACAA 801
 Db 227 TCTAACCACGAGTAAACAAATGCGTGAATGTGATGTAAGTGCACACTACAA 168
 Qy 802 CAGATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 861
 Db 167 CAGATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 108
 Qy 862 TTATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 921
 Db 107 TTATTCCTTACCGTCTCCACAAAGGTGAGATTTGTAATGTCATACCTTTT 48
 Qy 922 CTGCTTTATTAATAATGATGTGTATCTTGTCTTCTTCCGATACC 968
 Db 47 CTGCTTTATTAATAATGATGTGTATCTTGTCTTCCGATACC 1

RESULT 10
 AAI1818/C
 ID AAI1818 standard; DNA: 487 BP.

AC AAI1818;

DT 12-OCT-2001 (first entry)

DE Probe #1751 for gene expression analysis in human cervical cell sample.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

KM cervical cancer; ss.

OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for
 analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID NO 1751; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs). The present sequence is one such probe. The SENPs are derived
 CC from human Hela cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a

CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pat_sequences](http://wipo.int/pub/published_pat_sequences).

SQ Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match	33.2%	Score 467;	DB 22;	Length 487;
Best Local Similarity	100.0%	Pred. No. 2.4e-55;		
Matches 467;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	507	AGGAATTTGCGTAAATCCCTGGAAATTCCTTTTGGAAACGAGCTAAAGATGCAACA	561
Db	467	AGGAATTTGCGTAAATCCCTGGAAATTCCTTTTGGAAACGAGCTAAAGATGCAACA	408
Qy	562	ATGTAGAAGACGTCCTTCAATGACGATGGCCAGCTGAGATTAAACGAAATGGTCCGGAG	621
Db	407	ATGTAGAAGACGTCCTTCAATGACGATGGCCAGCTGAGATTAAACGAAATGGTCCGGAG	348
Qy	622	CACACGCTGGTGGTGGTGGAGTCCAAATGTTAAATTCACAGACATCCGACGACGAGT	681
Db	347	CACACGCTGGTGGTGGTGGAGTCCAAATGTTAAATTCACAGACATCCGACGACGAGT	288
Qy	682	CAGGTGGAGGTTCCTGCTAAATTTGGCTCCGATCCCTTTTCGACAGCAATGAATTGGAA	741
Db	287	CAGGTGGAGGTTCCTGCTAAATTTGGCTCCGATCCCTTTTCGACAGCAATGAATTGGAA	228
Qy	742	TCTGACCCCAAGTGAAGAAAACAAATTCGCTGAATGTATACGTATATAGCTGCATACAA	801
Db	227	TCTGACCCCAAGTGAAGAAAACAAATTTGCTGATATGTATATAGTATAGCTGCATACAA	168
Qy	802	CAGATCTCTACGCTCTCCACCAAGCTCAGAGATGTGAATATGGTAACTACTGCTTTTTT	861
Db	167	CAGATCTCTACGCTCTCCACCAAGCTCAGAGATGTGAATATGGTAACTACTGCTTTTTT	108
Qy	862	TTATTTCCCTTGACTCAGACGACCTACTTCATTTTCAGACGCTGTTTAAACCTTTGTGTG	921
Db	107	TTATTTCCCTTGACTCAGACGACCTACTTCATTTTCAGACGCTGTTTAAACCTTTGTGTG	48
Qy	922	CTGGTTTAAATTAATGTGTATATCTTGTGCTCTTCGTGATACC	968
Db	47	CTGGTTTAAATTAATGTGTATATCTTGTGCTCTTCGTGATACC	1

RESULT 11	AA133136//C	
ID	AA133136 standard; DNA; 487 Bp.	
XX		
AC	AA133136;	
XX		
DT	17-OCT-2001 (first entry)	
DE	Probe #1822 used to measure gene expression in human placenta sample.	
XX		
XX		
KW	Probe: microarray; human; placenta; antenatal diagnosis;	
KW	genetic disorder; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157272-A2.	
XX		
PD	09-AUG-2001.	
PE	30-JAN-2001; 2001WO-US000663.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
PR	03-AUG-2000; 2000US-0632366.	
PR	21-SEP-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236359.	

PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR
XX WPI; 2001-488897/53.

PT	Human genome-derived single exon nucleic acid probes useful for
PT	analyzing gene expression in human placenta -
XX	
Claim 25; SEQ ID NO 1822; 654bp; English.	
PS	

CC The present invention relates to single exon nucleic acid probes (SENAP)
CC
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.

50 Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match	33.2%	Score 467;	DB 22;	Length 487;
Best Local Similarity	100.0%	Pred. No. 2.4e-55;		
Matches 467; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

QY	502	AGAAATTTGCTGATTCCTCCGGAATTCCTGTTTGGAAACAGAGCTAGAGATGCAACGA	561
QY	503	AGAAATTTGCTGATTCCTCCGGAATTCCTGTTTGGAAACAGAGCTAGAGATGCAACGA	562
Db	467	AGAAATTTGCTGATTCCTCCGGAATTCCTGTTTGGAAACAGAGCTAGAGATGCAACGA	4080
QY	562	ATGTGAAACAGCTCTTCAATGACGATGGCAGCAGCAATTTAAAGACGAATGGGTCCGGAG	6211
Db	407	ATGTGAAACAGCTCTTCAATGACGATGGCAGCAGCAATTTAAAGACGAATGGGTCCGGAG	3488
QY	622	CAACAGCTGGTGGTGTGATGAAGTCGAATGTTAAATTCACAGACATCCAGCAAGCAAGT	6814
Db	347	CAACAGCTGGTGGTGTGATGAAGTCGAATGTTAAATTCACAGACATCCAGCAAGCAAGT	2888
QY	682	CAGGTGAGGTTCGTCGTAATTTGCTCCCAATCCTTTTCACACAGATGAAATTTGCA	7411
Db	287	CAGGTGAGGTTCGTCGTAATTTGCTCCCAATCCTTTTCACACAGATGAAATTTGCA	2288
QY	742	TCTGAACCCAAAGTAAAAACAAATTTGCTGATTTGACTGATGACTGACTACAA	8014
Db	227	TCTGAACCCAAAGTAAAAACAAATTTGCTGATTTGACTGATGACTGACTACAA	1688
QY	802	CAGATTTCTACCGTCTCCACAAAGTCGAGATCTGAATGGTCATTAAGTCACTTTT	8614
Db	167	CAGATTTCTACCGTCTCCACAAAGTCGAGATCTGAATGGTCATTAAGTCACTTTT	1080
QY	862	TTATTTCCCTTGACTCAAGACAGCTAACTTATTTTCAGACTGTTTAAACCTTTGTGTG	9211
Db	107	TTATTTCCCTTGACTCAAGACAGCTAACTTATTTTCAGACTGTTTAAACCTTTGTGTG	48
QY	922	CTGGTTTATTAATAATGATGATCTCTTTGGTCTCGATACG	968
Db	47	CTGGTTTATTAATAATGATGATCTCTTTGGTCTCGATACG	1

RESULT 12	
AA101752/c	
ID	AA101752 standard; DNA; 487 BP.
XX	
XX	
AC	AA101752:
XX	
XX	
DT	09-OCT-2001 (first entry)
XX	
DE	Probe #1743 used to measure gene expression in human breast sample.
XX	
XX	
XX	Probe: human; breast disease; breast cancer; development disorder; ss;
KW	inflammatory disease; proliferative breast disease; non-carcinoma tumour
XX	
OS	Homo sapiens.
XX	
XX	

PN W0200157270-A2.
 XX 09-AUG-2001.
 XX 29-JAN-2001; 2001WO-US00661.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-476286/51.
 XX Novel single exon nucleic acid probe used to measuring gene expression
 PT in a human breast -
 XX Claim 25; SEQ ID No 1743; 322pp; English.
 PS The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 CC
 XX Sequence 487 BP; 154 A; 97 C; 98 G; 138 T; 0 other;

Query Match 33.2%; Score 467; DB 22; Length 487;

Best Local Similarity 100.0%; Pred. No. 2,4e-55; Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 AGGAATTTGCGATTCCTTGGAAATCCGTTTGGAAACCAAGCTAAGAATGCAACGA 561
 DB 467 AGGAATTTGCGATTCCTTGGAAATCCGTTTGGAAACCAAGCTAAGAATGCAACGA 408
 QY 562 ATGTAGAACAGCTTTTCATGACGATGCGAGCTGAGATTAAAAAGCAATGGGTCGCGAG 621
 DB 407 ATGTAGAACAGCTTTTCATGACGATGCGAGCTGAGATTAAAAAGCAATGGGTCGCGAG 348
 QY 622 CAACAGCGGTGGTGGTGAAGAGTCCATGTTAAATCAAGACCTCCATGCAAGCA 681
 DB 347 CAACAGCGGTGGTGGTGAAGAGTCCATGTTAAATCAAGACCTCCATGCAAGCA 288
 QY 682 CAGGTGAGGTGGTGGTGAAGATTTGCTCCATCTTTTCACAGCAATGAATTTGCAA 741
 DB 287 CAGGTGAGGTGGTGGTGAAGATTTGCTCCATCTTTTCACAGCAATGAATTTGCAA 228
 QY 742 TCTGAACCAAGTGAAGAAACAAATTCCTGAATTTGACTGTATGTAGCTGCACATACAA 801
 DB 227 TCTGAACCAAGTGAAGAAACAAATTCCTGAATTTGACTGTATGTAGCTGCACATACAA 168
 QY 802 CAGATTTTACCGTCTCCACAAAGTGCAGATTTGTAATGTCAATCTGACTTTT 861
 DB 167 CAGATTTTACCGTCTCCACAAAGTGCAGATTTGTAATGTCAATCTGACTTTT 108
 QY 862 TTATTTCCCTTACCTCAACAGAGCTAATTTTCAAGAGCTGTTTAAACCTTTGTGTG 921
 DB 107 TTATTTCCCTTACCTCAACAGAGCTAATTTTCAAGAGCTGTTTAAACCTTTGTGTG 48

QY 922 CTGTTTATTAATAATGTGTGTAATCCTTGTGCTTTCGTATACC 968
 DB 47 CTGTTTATTAATAATGTGTGTAATCCTTGTGCTTTCGTATACC 1
 RESULT 13
 ID ABS01772/c
 ID ABS01772 standard; DNA; 487 BP.
 AC ABS01772;
 XX 19-AUG-2002 (first entry)
 XX Human genome-derived single exon probe from lung SEQ ID No 1763.
 DE Human; ds; single exon probe; asthma; lung cancer; COPD; IID;
 KW Chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS W0200186003-A2.
 XX 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US00665.
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 1; SEQ ID No 1763; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of

DT 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 14052.
DE
XX
XX Human: prostate cancer: cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX
PS Claim 1; Page 2343-2344; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX
SQ Sequence 406 BP; 104 A; 99 C; 92 G; 111 T; 0 other;
XX
XX
Query Match 18.8%; Score 264; DB 23; Length 406;
Best Local Similarity 100.0%; Pred. No. 9.7e-28;
Matches 264; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 518 CCTGGAATTCGTTTGAACAGCTAGAAATGCAAGCAATGTAGAACTCTT 577
DB 406 CCTGGAATTCGTTTGAACAGCTAGAAATGCAAGCAATGTAGAACTCTT 347
OY 578 CATGACATGCGAGCTGAGATTAAAGCAATGAGTCCCGAGCAACAGCTGTGTGC 637
DB 346 CATGACATGCGAGCTGAGATTAAAGCAATGAGTCCCGAGCAACAGCTGTGTGC 287
OY 638 TGAGAGTCGAATGTTAAATTCAGAGCACTCCAGTCAAGCAAGTGAAGTTGCTG 697
DB 286 TGAGAGTCGAATGTTAAATTCAGAGCACTCCAGTCAAGCAAGTGAAGTTGCTG 227
OY 698 CTAATAATTCGCTCCATCTTTCACAGCAATGAATTCGAATCTGAACCAAGTGAA 757
DB 226 CTAATAATTCGCTCCATCTTTCACAGCAATGAATTCGAATCTGAACCAAGTGAA 167
OY 758 AAACAAATTCGCTGAATGTAC 781

DB ||||||||||||||||||||
166 AAACAAATTCGCTGAATGTAC 143

Search completed: November 17, 2002, 11:42:33
Job time : 261 secs

Dh 241 GGGTTGGAAAGCTTCCCTCTCTAGTTGAGATGATACATATACAGAAAGCTACA 300
Qy 301 TCACACAAATGTGTGATTTCAAAATAGAACTATAGAGTTAGACGGGAAACAACTCA 360
Dh 301 TCACACAAATGTGTGATTTCAAAATAGAACTATAGAGTTAGACGGGAAACAACTCA 360
Qy 361 AGCTTCAAAATAGAGTTCTTCAATATGTTAAACAGTGGCTCAGAGAAATAGCTTATG 420
Dh 361 AGCTTCAAAATAGAGTTCTTCAATATGTTAAACAGTGGCTCAGAGAAATAGCTTATG 420
Qy 421 CCAGTGAATAATGTCAACAAATGTGTGAGGAGCAAAATGATCTGACCAAAAGAAAG 480
Dh 421 CCAGTGAATAATGTCAACAAATGTGTGAGGAGCAAAATGATCTGACCAAAAGAAAG 480
Qy 481 TAGTAGACTACACAACAGCGAAGAAATTTGCTGATTCCTTGGAAATCCGTTTGGAAA 540
Dh 481 TAGTAGACTACACAACAGCGAAGAAATTTGCTGATTCCTTGGAAATCCGTTTGGAAA 540
Qy 541 CCAAGTGTAGAAATGCAACGAATGTAGAACAGTCTTTCATGACGATGCGACCTGAGATTA 600
Dh 541 CCAAGTGTAGAAATGCAACGAATGTAGAACAGTCTTTCATGACGATGCGACCTGAGATTA 600
Qy 601 AAAAGCGAATGGGTCCCGAGACACAGCTGGTGGTCTGAGAAATGCCAATGTTAAATTC 660
Dh 601 AAAAGCGAATGGGTCCCGAGACACAGCTGGTGGTCTGAGAAATGCCAATGTTAAATTC 660
Qy 661 AGAGCACTCCAGTCAAGCAGTCAAGTGGAGGTGCTGCTAAATTTGCTCATCTCTT 720
Dh 661 AGAGCACTCCAGTCAAGCAGTCAAGTGGAGGTGCTGCTAAATTTGCTCATCTCTT 720
Qy 721 CTCACAGCAATGAAATTTGCAATCTGAACCCAGTGAAGAAACAAATTTGCTCATCTT 780
Dh 721 CTCACAGCAATGAAATTTGCAATCTGAACCCAGTGAAGAAACAAATTTGCTCATCTT 780
Qy 781 CTGATAGTAGTGGACATCAACAGATCTTACCGTCCCAAAAGGTGAGATTTGTA 840
Dh 781 CTGATAGTAGTGGACATCAACAGATCTTACCGTCCCAAAAGGTGAGATTTGTA 840
Qy 841 TGGTCAATACTGACTTTTTTTTATTCCTTCCCTGACTCAAGACAGTAACTTATTTTCA 900
Dh 841 TGGTCAATACTGACTTTTTTTTATTCCTTCCCTGACTCAAGACAGTAACTTATTTTCA 900
Qy 901 ACTGTTTAAACCTTTTGTGCTGCTGTTATATAAATATGTGTATCTTGTGCTTTC 960
Dh 901 ACTGTTTAAACCTTTTGTGCTGCTGTTATATAAATATGTGTATCTTGTGCTTTC 960
Qy 961 CTGATACAGCTGTTTCCCGTGGTGGTGAATATTTTGTGATGATTTATATG 1020
Dh 961 CTGATACAGCTGTTTCCCGTGGTGGTGAATATTTTGTGATGATTTATATG 1020
Qy 1021 GCATGTTAGATGTCAGGTTAGTCTTCTGAAGATGAAGTTCAGCCATTTTGTATCAAC 1080
Dh 1021 GCATGTTAGATGTCAGGTTAGTCTTCTGAAGATGAAGTTCAGCCATTTTGTATCAAC 1080
Qy 1081 AGCAACAGCAGTGTCTGCTCACTTCCATGATTAAGTTTATGAGATTTATATGTA 1140
Dh 1081 AGCAACAGCAGTGTCTGCTCACTTCCATGATTAAGTTTATGAGATTTATATGTA 1140
Qy 1141 TCTATATTTGCTAGTCTTCTCTGTAGAGTTAAATGGAAGATTTACCTTCTGATTTA 1200
Dh 1141 TCTATATTTGCTAGTCTTCTCTGTGTAGAGTTAAATGGAAGATTTACCTTCTGATTTA 1200
Qy 1201 TAGTTTCTTCACTGATGATATTTGCTGAGAAATTTGTAATTTGTTGACAC 1260
Dh 1201 TAGTTTCTTCACTGATGATATTTGCTGAGAAATTTGTAATTTGTTGACAC 1260
Qy 1261 TATGTAAACAAACAGTGAAGATATGTTTAAATATGTATGTAATTTGGAAGTAAAA 1320
Dh 1261 TATGTAAACAAACAGTGAAGATATGTTTAAATATGTATGTAATTTGGAAGTAAAA 1320
Qy 1321 AAA 1380
Dh 1321 AAA 1380

Qy 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405
Dh 1381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405

RESULT 2
US-09-820-003a-3
Sequence 3, Application US/09820003A
Patent No. US20020142382A1
GENERAL INFORMATION:
APPLICANT: MERKULOV, Gennady et al.
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: CL001196
CURRENT APPLICATION NUMBER: US/09/820,003A
CURRENT FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 46050
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(46050)
OTHER INFORMATION: n = A,T,C or G
US-09-820-003a-3

Query Match 58.28; Score 817.2; DB 10; Length 46050;
Best Local Similarity 99.6%; Pred. No. 2,7e-105;
Matches 819; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 502 AGAATTTGCTGATTCCTTCCCTGGAATTCCTTTTGGAAACAGGTCMAAATGCAACGA 561
Dh 43238 AGAATTTGCTGATTCCTTCCCTGGAATTCCTTTTGGAAACAGGTCMAAATGCAACGA 43297
Qy 562 ATGTAGAACAAGCTTTTCATGACGATGACGAGTGAATTTAAACGGAATGGGTCCGGAG 621
Dh 43298 ATGTAGAACAAGCTTTTCATGACGATGACGAGTGAATTTAAACGGAATGGGTCCGGAG 43357
Qy 622 CAACAGCTGTGTGTGTGAGAGTCCAAATGTTAAATTCAGAGCAGTCCAGTCAAGCAGT 681
Dh 43358 CAACAGCTGTGTGTGTGAGAGTCCAAATGTTAAATTCAGAGCAGTCCAGTCAAGCAGT 43417
Qy 682 CAGGTGAGAGTGTGCTGTAATTTGCTCCATCCCTTTTTCACAGCAATGAATTTGCCA 741
Dh 43418 CAGGTGAGAGTGTGCTGTAATTTGCTCCATCCCTTTTTCACAGCAATGAATTTGCCA 43477
Qy 742 TCTGAACCCAAAGTGAATAATGCTGAATTTGCTGATGATGATGATGATGATGATGAT 801
Dh 43478 TCTGAACCCAAAGTGAATAATGCTGAATTTGCTGATGATGATGATGATGATGATGAT 43537
Qy 802 CAGATTTCTTACGCTCCACCAAGGTGAGAGTGAATTTGCTGATGATGATGATGATGAT 861
Dh 43538 CAGATTTCTTACGCTCCACCAAGGTGAGAGTGAATTTGCTGATGATGATGATGATGAT 43597
Qy 862 TTATTTCCCTGACACAGACAGTAACTTCAATTTTCAAGAACTGTTTAAACCTTTGCTG 921
Dh 43598 TTATTTCCCTGACACAGACAGTAACTTCAATTTTCAAGAACTGTTTAAACCTTTGCTG 43657
Qy 922 CTGTTTATATAAATATGTTGTAATCTTGTGCTTCCGATGACAGAGTGTTCGCG 981
Dh 43658 CTGTTTATATAAATATGTTGTAATCTTGTGCTTCCGATGACAGAGTGTTCGCG 43717
Qy 982 TGGTTGTTAGAAATATATTTTGTGTAATTTATTTGATGATTTAGATGTCAGGTTT 1041
Dh 43718 TGGTTGTTAGAAATATATTTTGTGTAATTTATTTGATGATTTAGATGTCAGGTTT 43777
Qy 1042 AGCTTTCTGAAGATGAAGTTCAGCATTGTTATCAAAACAGACACAGCAGTGTGTGCAC 1101
Dh 43778 AGCTTTCTGAAGATGAAGTTCAGCATTGTTATCAAAACAGCAGCAGTGTGTGCAC 43837

Qy	1102	TTTCACGCACTAAATTAGAGAGAGTTATATGTAAGATCGATTGCGATTCCTCT	1161
Db	43838	TTTCCATCGCACTAAATTAGAGAGAGTTATATGTAAGATCGATTGCGATTCCTCT	43897
Qy	1162	TTTAGAGTTATAAATCGAAAGATTACATCTGATTAATAGTTCTTCATCTCTGAT	1221
Db	43899	TGTAGAGTTATAAATCGAAAGATTACATCTGATTAATAGTTCTTCATCTCTGAT	43959
Qy	1222	ATAATTTGGCGTGCAGAAATATTTGTATATTTGGCGACATATGTAACAAACAATGAAG	1281
Db	43959	ATAATTTGGCGTGCAGAAATATTTGTATATTTGGCGACATATGTAACAAACAACCTGAAG	44017
Qy	1282	ATATGTTTAATTAATATTGTACTATTGGAGATTAATAAAAAA	1323
Db	44018	ATATGTTTAATTAATATTGTACTATTGGAGATTAATTAATAA	44059

RESULT 3

US-09-820-003A-37
Sequence 37, Application US/09820003A
Patent No. US20020142382A1
GENERAL INFORMATION:
APPLICANT: MERKULOV, Gennady et al.
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE PROTEINS,
TITLE OF INVENTION: PROTEINS AND USES THEREOF
FILE REFERENCE: CL001196
CURRENT APPLICATION NUMBER: US/09/820, 003A
CURRENT FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 37
LENGTH: 601
TYPE: DNA
ORGANISM: Homo sapien
US-09-820-003A-37

	Query Match	Best Local Similarity	92.08 %	Score 589.8	DB 10	Length 601	
	Matches 591	Conservative 1		Mismatches 3	Indels 0	Gaps 0	
QY	729	AATGATTTTCACATGTGACACCAAGGAAAAAACAATTTGGCTCAATTTACTGTATGT					788
Db	1	AATGATTTTCACATGTGACACCAAGGAAAAAACAATTTGGCTCAATTTACTGTATGT					60
QY	789	AGCGGACATCAACAGATTTCTTTCGCTCCGTCGACAAAGTCACAGATTCCTAAAGCGCAT					848
Db	61	AGCGGACATCAACAGATTTCTTTCGCTCCGTCGACAAAGTCACAGATTCCTAAAGCGCAT					120
QY	849	ACTACACTTTTTTTTATTTCCCTCGATGACACAGCGTAACCTCATTTCCAGACAGCTTTT					908
Db	121	ACTACACTTTTTTTTATTTCCCTCGATGACACAGCGTAACCTCATTTCCAGACAGCTTTT					180
QY	909	AAACCTTGGGCGCGGTTTATAAATATGTGTATATTCCTGTGCTTCTCTATACC					968
Db	181	AAACCTTGGGCGCGGTTTATAAATATGTGTATATTCCTGTGCTTCTCTATACC					240
QY	969	AGACTGTTTCCCGTGGTGGTATATATTTTGTGTATGTATATATGGCATGTTT					1028
Db	241	AGACTGTTTCCCGTGGTGGTATATATTTTGTGTATGTATATATATGGCATGTTT					300
QY	1029	AGATGTCAGGTTTACGTCTTCTGAAGTGAAGTTCAGCCATTTTGTATCAACAGCAAG					1088
Db	301	AGATGTCAGGTTTACGTCTTCTGAAGTGAAGTTCAGCCATTTTGTATCAACAGCAAG					360
QY	1089	CAGTGTCTGTCACTTTCCATGACATATAAGTTTAGAGATGATATATGTAAAGATCTGATT					1148
Db	361	CAGTGTCTGTCACTTTCCATGACATATAAGTTTAGAGATGATATATGTAAAGATCTGATT					420
QY	1149	GCTAGGTTCTTCCCTTGTAGAGATTATAATGTGAAGAAATTAACACTATCTGATTAATAGTTTCT					1208
Db	421	GCTAGGTTCTTCCCTTGTAGAGATTATAATGTGAAGAAATTAACACTATCTGATTAATAGTTTCT					480

Qy	1209	TCTACCTCTCATATTAATTTGGCTGCAGAAATTTATTTGGTGCACATCTAGAC	1268
Db	481	TCTATCTCTCATATTAATTTGGCTGCAGAAATTTATTTGGTGCACATCTAGAC	540
Qy	1269	AAAAACATGAGATATCTTTAATAAATTTGTACTATTCGACGATAAAAAAA	1323
Db	541	AAAAACATGAGATATCTTTAATAAATTTGTACTATTCGACGATATATCAAAA	595

RESULT 4

US-09-920-300A-374/C
Sequence 374, Application US/09920300A
Patent No. US20020136728A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeline Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
OF COLON CANCER
FILE REFERENCE: 210121.547
CURRENT APPLICATION NUMBER: US/09/920.300A
CURRENT FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 1789
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 374
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
US-09-920-300A-374

	Query Match	Score	DB 10%	Length
	Best Local Similarity	100.0%	Pred. No. 2.3e-63	506
	Matches	Conservative	0	Mismatches 0; Indels 0; Gaps 0
QY	770	CCGAAATGCTAGCTATGCTAGCGACATCAAGAGATCTCTCCGCTCCGACAAAAGCA	829	
Db	506	CCGAAATGCTAGCTATGCTAGCGACATCAAGAGATCTCTCCGCTCCGACAAAAGCA	447	
QY	830	GAGATGTAAAGGCAATACGACATCTTTTATCCCTGACACGACGCACT	889	
Db	446	GAGATGTAAAGGCAATACGACATCTTTTATCCCTGACACGACGCACT	387	
QY	890	TGATTTTGACAACTCTTTTAAACCTTTGGGCGCGGTATTAATAATTAATGTGTATTC	949	
Db	386	TCATTTTGACAACTCTTTTAAACCTTTGGGCGCGGTATTAATAATTAATGTGTATTC	327	
QY	950	TTGTGTGCTTCTCTGATACACAGACTGTTCCGCTGGTGGATATATTTGTTTGA	1009	
Db	326	TTGTGTGCTTCTCTGATACACAGACTGTTCCGCTGGTGGATATATTTGTTTGA	267	
QY	1010	TGTTTATATTCGCATGTTTGAGATGACAGTTTAACTCTCTGAAGATGAAGTTAGGCATT	1069	
Db	266	TGTTTATATTCGCATGTTTGAGATGACAGTTTAACTCTCTGAAGATGAAGTTAGGCATT	207	
QY	1070	TTGTATCAACACGACACAGAGGTGCTGTCACTTCGATGCAATAAAGTTTACTGATGATGT	1129	
Db	206	TTGTATCAACACGACACAGAGGTGCTGTCACTTCGATGCAATAAAGTTTACTGATGATGT	147	
QY	1130	TATATGTAAAGATCTGATTTGCTATAGTCTTCCTGTGAGCTATTAATAATGAGAAATTAAC	1189	
Db	146	TATATGTAAAGATCTGATTTGCTATAGTCTTCCTGTGAGCTATTAATAATGAGAAATTAAC	87	
QY	1190	TATCTGAATTAAGTTCTTCTCAATCTGCATATATAATTTTGCGCTGAGAAATTTGTAAT	1249	
Db	86	TATCTGAATTAAGTTCTTCTCAATCTGCATATATAATTTTGCGCTGAGAAATTTGTAAT	27	
QY	1250	TTGTTGACACTATGTAAACAAACA	1275	
Db	26	TTGTTGACACTATGTAAACAAACA	1	

RESULT 5
US-10-033-528-374/c
; Sequence 374, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY.
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 374
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-374

Query Match 36.0%; Score 506; DB 12; Length 506;
Best Local Similarity 100.0%; Pred. No. 2.3e-62;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 770 CCTGAATGTACTGTATGTAGTGCACATCAAGATTCCTACGCTCCACAAGGTCA 829
DB 506 CCTGAATGTACTGTATGTAGTGCACATCAAGATTCCTACGCTCCACAAGGTCA 447
QY 830 GAGATTTGAATGGTCAATCTACTGACTTTTATTCCTGAGCTCAAGACGTAAT 889
DB 446 GAGATTTGAATGGTCAATCTACTGACTTTTATTCCTGAGCTCAAGACGTAAT 387
QY 890 TCATTTAGAACGCTTTTAAACCTTGTGCTGCTTTATTAATAATATGTTGAATCC 949
DB 386 TCATTTAGAACGCTTTTAAACCTTGTGCTGCTTTATTAATAATATGTTGAATCC 327
QY 950 TTGTTGCTTCTGTATACAGACTGTTCCCGTGGTGGTGAATATATTTTGTGTA 1009
DB 326 TTGTTGCTTCTGTATACAGACTGTTCCCGTGGTGGTGAATATATTTTGTGTA 267
QY 1010 TGTATTATTTGGCATGTTAATGTCAGGTTAGTCTTCTGGAAGTGAAGTTCA 1069
DB 266 TGTATTATTTGGCATGTTAATGTCAGGTTAGTCTTCTGGAAGTGAAGTTCA 207
QY 1070 TTGTATCAACAGACAGACAGTGTCTCACTTCCATGCAATAGTTAGTGAGAT 1129
DB 206 TTGTATCAACAGACAGACAGTGTCTCACTTCCATGCAATAGTTAGTGAGAT 147
QY 1130 TATATGTAAGATCTGATTTGCTAGTCTTCTGTAGAGATTATAAATGGAAGATTAC 1189
DB 146 TATATGTAAGATCTGATTTGCTAGTCTTCTGTAGAGATTATAAATGGAAGATTAC 87
QY 1190 TATCGATTAATAGTTTCTTCACTGCAATATATTTGGGCTGCAAGATTTGTA 1249
DB 86 TATCGATTAATAGTTTCTTCACTGCAATATATTTGGGCTGCAAGATTTGTA 27
QY 1250 TTGTTGCACACATGTATCAAAACA 1275
DB 26 TTGTTGCACACATGTATCAAAACA 1

RESULT 6
US-09-864-761-1742/c
; Sequence 1742, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 1742
;; LENGTH: 487
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC007318.2
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.6
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.4
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 10
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
US-09-864-761-1742
Query Match 33.2%; Score 467; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 5.6e-57;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 AGGAATTTGCTGATTCCTTGGAATTCGTTTGGAAACGAGCTAAGATGCAACGA 561
DB 467 AGGAATTTGCTGATTCCTTGGAATTCGTTTGGAAACGAGCTAAGATGCAACGA 408
QY 562 ATGTAGAAGACTTTTCAATGACGATGCGAGCTGAGATTAAACGAGATGGTCCGAG 621
DB 407 ATGTAGAAGACTTTTCAATGACGATGCGAGCTGAGATTAAACGAGATGGTCCGAG 348
QY 622 CAACAGCTGTGTGCTGAGAGATCCATGTTAAATTCAGAGCAGCTCCAGCAGT 681


```

: FEATURE: misc:feature
: NAME/KEY: misc:feature
: OTHER INFORMATION: Incyte ID No. US20010051335A1 700354602H1
: NAME/KEY: unsure
: LOCATION: 2, 88, 216, 219, 247, 283
: OTHER INFORMATION: a, t, c, g, or other
: US-09-294-093B-4584

```

Query Match	18.6%	Score 261	DB 10	Length 290
Best Local Similarity	97.3%	Pred. No. 1.6e-28		
Matches 283	Conservative 0	Mismatches 6	Indels 2	Gaps 2

Oy	540	ACAGTGGCTAAGAATTCGCAACAATGTGAACAGCTTTCAAGACAGTGGCAGATG	595
Dd	1	AMCACTGCTTAAGAATGCACACAAATGTGAACAGCTTTCAATGACAGTAGGCCAGCTGAATG	60
Oy	600	AAAAAGCGAATGGGTCCCGGAGCAACAGCTGGTGGT-GCTGAGAGATCCATGTTAAAT	658
Dd	61	AAAAAGCGAATGGGTCCCGGAGCAACACTGGTGGTGCTGAGAGATCCATGTTAAAT	120
Oy	659	TCAAGCACTCCAGTCAACAGCATGAGGAGAGTGGTGGCTAAATTTGGCTCCATCCTT	718
Dd	121	TCAGGCACTCCAGTCAACAGCATGAGGAGAGTGGTGGCTAAATTTGGCTCCATCCTT	180
Oy	719	TTCTCACAGCAATGATTTGCAATCTGAACCCAGTGAAAAACAAATTGCCCTGAATG	778
Dd	181	TTCTCACAGCAATGATTTGCAATCTGAACCCAGTAGNANNAACAAATTGCCCTGAATG	248
Oy	779	TACTGTATGTAGTCCTATAACAAGATTTACCCTGCTCCACAAGGTCA	829.
Dd	241	TACTGTATGTAGTC-GCACCTAACACAGATTTCTACCGTCTCCACAAAGGTCA	290

```

RESULT 10
US-09-604-287A-337
; Sequence 337, Application US/09604287A
; Patent No. US20020064872A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C7
CURRENT APPLICATION NUMBER: US/09/604,287A
CURRENT FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 337
LENGTH: 241
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(241)
OTHER INFORMATION: n = A,T,C or G
US-09-604-287A-337

```

	Query Match	16.6%	Score 232.8	DB 10	Length 241
	Best Local Similarity	97.5%	Pred. No. 1.3e-24		
	Matches 234	Conservative	0	Mismatches 6	Indels 0
				Gaps	0
QY	778	GTACGTATGTATGCTGCACATACACAGATTCTTACCGTCTCCACAAAGTGCAGATTGT	837		
Db	2	GTACGTATGTATGCTGCACATACACAGATTCTTACCGTCTCCACAAAGTGCATATTGT	61		
QY	838	AAATGTCATATCTGCATCTTTTTTATTCCTTGACCTCAAGCAGCTAACTTCATTTC	897		
Db	62	AAATGTCATATCTGCATCTTTTTTATTCCTTGACCTCAAGCAGCTAACTTCATTTC	121		

Oy	AGAATGTTTAAACCTTTGTCCTGGTTAAATAATATGTTATCCCTGTGGCT	957
898	AGAATGTTTAAACCTTTGTCCTGGTTAAATAATATGTTATCCCTGTGGCT	
Db	122 AGAATGTTTAAACCTTTGTCCTGGTTAAATAATATGTTATCCCTGTGGCT	181
Oy	958 TTTCGATACACAGCTGTTTCCGTCGTTCGTAGAAATATATTTGTTTGATGTTTATA	1011
Db	182 TTTCGATACACAGCTGTTTCCGTCGTTCGTAGAAATATATTTGTTTGATGTTTATA	241

```

RESULT 11
US-10-007-805-337
; Sequence 337, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margaila
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 337
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 47..56, 69..228
; OTHER INFORMATION: n = A,T,C or G
US-10-007-805-337

```

Query Match	16.6%	Score 232.8	DB 12	Length 241
Best Local Similarity	97.5%	Pred. No. 1.3e-24		
Matches 234	Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY	778	GTACTGTATGTAGCTGCACATACAGATCTTACCGTCTCCACANAAGGTGACAGATTGT	837	
Db	2	GTACTGTATGTAGCTGCACATACAGAGATTCTTACCGTCTCCACANAAGGTGACAGATTGT	61	
QY	838	AAATGTCATCTACATTTTATTCCTGACCTACACAGAGCTAACTCATTTTC	897	
Db	62	AAATGTCATCTACATTTTATTCCTGACCTACACAGAGCTAACTCATTTTC	121	
QY	898	AGAACTGTTTTAAACCTTGTGTGCTGTTTAAATTAATGTGHTAATCCCTGTGTCT	957	
Db	122	AGAACTGTTTTAAACCTTGTGTGCTGTTTAAATTAATGTGHTAATCCCTGTGTCT	181	
QY	958	TTTCCTGATACCAAGACTGTTCCCGTGGTGGTTAAGATATTTTGTGATGTTTAA	1017	
Db	182	TTTCCTGATACCAAGACTGTTCCCGTGGTGGTTAAGATATTTTGTGATGTTTAA	241	

```

RESULT 12
US-09-820-003A--36
; Sequence 36, Application US/09820003A
; Patent No. US20020142382A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01196

```

```

; CURRENT APPLICATION NUMBER: US/09/820,003A
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-820-003A-36

```

```

Query Match          15.9%; Score 224; DB 10; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.1e-23;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 502 AGGATTTGCTGATTCCTCGATTCGTTTGGAAACCGAGTGAAGATGCAACGA 561
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 378 AGGATTTGCTGATTCCTCGATTCGTTTGGAAACCGAGTGAAGATGCAACGA 437
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 562 ATGTAGAACAGCTCTTCATGACGATGGCAGTGAATTAACCAATGGGTCCCGAG 621
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 438 ATGTAGAACAGCTCTTCATGACGATGGCAGTGAATTAACCAATGGGTCCCGAG 497
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 622 CAACAGCTGGTGGTGTGAGAACGTCGAATGTAAATTCAGACGCTCCAGCAGCAGT 681
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 498 CAACAGCTGGTGGTGTGAGAACGTCGAATGTAAATTCAGACGCTCCAGCAGCAGT 557
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 682 CAGGTGAGGTCTCTGCTTAATTTGCCCTTCATCTTTTCTCAG 725
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 558 CAGGTGAGGTCTCTGCTTAATTTGCCCTTCATCTTTTCTCAG 601
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 13

```

US-09-778-320-84
; Sequence 84, Application US/09778320
; Patent No. US20010034052A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Day, Craig H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Wang, Tongtong
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.491C5
; CURRENT APPLICATION NUMBER: US/09/778,320
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: FastSeq for Windows Version 3.0
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapien
; NAME/KEY: misc_feature
; LOCATION: (1) (229)
; OTHER INFORMATION: n = A,T,C or G
US-09-778-320-84

```

```

Query Match          15.6%; Score 219.8; DB 10; Length 229;
Best Local Similarity 96.5%; Pred. No. 7.9e-23;
Matches 221; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 780 ACTGATAGTACGCTACACACAGATTTCTACCGCTCCCAAGGTCAGATTTGTA 839
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 ACTGATAGTACGCTACACACAGATTTCTACCGCTCCCAAGGTCAGATTTGTA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 840 ATGTCAATACGACTTTTTCCTTCCCTGACGTCAGACGCTTAACCTATTTTTCAG 899
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ATGTCAATACGACTTTTTCCTTCCCTGACGTCAGACGCTTAACCTATTTTTCAG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 900 AACTGTTTAAACCTTTTGTGTGTTTATAAATAATATGTGTATCTTGTCTTT 959
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

DB 121 AACTGTTTAAACCTTTTGTGTGTTTATAAATAATATGTGTATCTTGTCTTT 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 960 CCGATACACAGACTGTTTCCCGTGGTGTAGATATATTTTGTGTTT 1008
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 CCGATACACAGACTGTTTCCCGTGGTGTAGATATATTTTGTGTTT 229
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14

```

US-09-910-689-84
; Sequence 84, Application US/09910689
; Patent No. US20020081609A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Day, Craig H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Wang, Tongtong
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.491C6
; CURRENT APPLICATION NUMBER: US/09/910,689
; CURRENT FILING DATE: 2001-07-20
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 163, 191, 203, 222, 223, 228
; OTHER INFORMATION: n = A,T,C or G
US-09-910-689-84

```

```

Query Match          15.6%; Score 219.8; DB 10; Length 229;
Best Local Similarity 96.5%; Pred. No. 7.9e-23;
Matches 221; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 780 ACTGATAGTACGCTACACACAGATTTCTACCGCTCCCAAGGTCAGATTTGTA 839
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 ACTGATAGTACGCTACACACAGATTTCTACCGCTCCCAAGGTCAGATTTGTA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 840 ATGTCAATACGACTTTTTCCTTCCCTGACGTCAGACGCTTAACCTATTTTTCAG 899
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ATGTCAATACGACTTTTTCCTTCCCTGACGTCAGACGCTTAACCTATTTTTCAG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 900 AACTGTTTAAACCTTTTGTGTGTTTATAAATAATATGTGTATCTTGTGTTT 959
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 AACTGTTTAAACCTTTTGTGTGTTTATAAATAATATGTGTATCTTGTGTTT 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 960 CCGATACACAGACTGTTTCCCGTGGTGTAGATATATTTTGTGTTT 1008
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 CCGATACACAGACTGTTTCCCGTGGTGTAGATATATTTTGTGTTT 229
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 15

```

US-10-010-742-84
; Sequence 84, Application US/10010742
; Patent No. US20020146727A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Day, Craig H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Wang, Tongtong
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Bennington, Angela Ann

```

```

; APPLICANT: Zehentner, Barbara
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.491C7
; CURRENT APPLICATION NUMBER: US/10/010,742
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 163, 191, 203, 222, 223, 228
; OTHER INFORMATION: n = A,T,C or G
US-10-010-742-84

```

```

Query Match          15.6%; Score 219.8; DB 12; Length 229;
Best Local Similarity 96.5%; Pred. No. 7.9e-23;
Matches 221; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 780 ACTGTATGAGTGCACACAGATTTTACCGTCTCCACAAGGTGAGATTGTAA 839
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 ACTGTGTAGTGCACACAGATTTTACCGTCTCCACAAGGTGAGATTGTAA 60

QY 840 ATGTCATATCTGACTTTTATTATCCCTGACTCAGACAGCTAATTCATTTCAG 899
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 ATGTCATATCTGACTTTTATTATCCCTGACTCAGACAGCTAATTCATTTCAG 120

QY 900 AACTGTTTAAACCTTTGTGCTGCTTTTAAATAATGTTAATCCTTTGCTTT 959
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 AACTGTTTAAACCTTTGTGCTGCTTTTAAATAATGTTAATCCTTTGCTTT 180

QY 960 CCTGATACCAAGCTGTTCCCGTGGTGTAGAAATATATTGTTTGTG 1008
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 CCTGATACCAAGCTGTTCCCGTGGTGTAGAAATATATTGTTTGTG 229

```

Search completed: November 17, 2002, 13:04:08
 Job time : 132 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 11:36:02 ; Search time 1933 Seconds
(without alignments)
11771.681 Million cell updates/sec

Title: US-09-820-003a-1

Perfect score: 1405
Sequence: 1 aagcagatagcgtgctgcgc.....aaaaaaaaaaaaaaaa 1405

Scoring table: IDENTITY_MNC
Gapop 10.0 ; Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST: *
1: em_estdb: *
2: em_esthum: *
3: em_estlin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_tod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	743	52.9	948	9	AL539022
2	732.4	52.1	777	14	B0014597
3	727.6	51.8	811	9	AA746643
4	723.8	51.5	817	13	B1869965
5	708.8	50.4	910	14	B0276678
6	708.4	50.4	761	14	B0772048

7	707.8	50.4	945	9	AL530265
8	704.6	50.1	1105	13	BM452262
9	703.6	50.1	725	14	B0448090
10	702	50.0	796	13	B1913092
11	692.8	49.3	760	14	BM971301
12	691.8	49.2	976	14	BM919860
13	690.6	49.2	837	13	B1086445
14	689	49.0	837	13	B1086445
15	688.2	49.0	689	14	BM706159
16	686	48.8	737	13	B1092003
17	685.6	48.8	686	14	BM826571
18	683.2	48.6	717	9	AI95257
19	682.6	48.6	706	14	BM975245
20	677.6	48.2	877	12	B1659520
21	676.4	48.1	678	13	BG939358
22	671.2	47.8	684	13	B1495589
23	670.8	47.7	1130	13	BM542833
24	669.6	47.6	813	14	B0178972
25	668.2	47.6	812	14	B0178972
26	661.2	47.1	686	14	B0599300
27	659.6	46.9	721	9	AI884351
28	658	46.8	793	12	B0178268
29	657.6	46.8	680	10	AM150848
30	655.2	46.5	682	14	B0010358
31	650.4	46.3	693	9	AI829874
32	650	46.3	908	9	AM559410
33	648	46.0	723	9	AI800097
34	645.6	46.0	682	9	AI858694
35	640.4	45.6	900	12	B0032817
36	638.8	45.5	682	9	AI860577
37	638	45.4	638	14	BM783973
38	638	45.4	646	14	B0599347
39	638	45.4	667	14	BE646347
40	637.4	45.4	719	14	B0210709
41	632.8	45.0	1055	13	BM472577
42	631.8	44.9	664	10	BE617921
43	631.2	44.9	645	14	B0232273
44	630.6	44.9	645	14	AM166902
45	628.6	44.7	702	9	AI889183

ALIGNMENTS

RESULT 1
AL539022
LOCUS AL539022 LTI_F1013.FBn1 Homo sapiens cDNA clone CSDB030YG06 5
DEFINITION AL539022 LTI_F1013.FBn1 Homo sapiens cDNA clone CSDB030YG06 5
ACCESSION AL539022 GI:12867866
VERSION AL539022.1 GI:12867866
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
location/Qualifiers
1. 948
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDB030YG06"
/clone_lib="LTI_F1013.FBn1"
/dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"
/lab_host="DH10B"

QY 613 GTCGGGAGCAACAGCTGGTGGCTGAGAGTCGCAATTTAAATTCAGACACTCCAG 672
 DB 718 GTCCGGGAGCAACAGCTGGTGGCTGAGAGTCGCAATTTAAATTCAGACACTCCAG 659
 QY 673 TCAGACAGTCAGGTGAGGTGCTGCTAAATTTGGCTCCATCTTTTCTCAGACGAATG 732
 DB 658 TCAGACAGTCAGGTGAGGTGCTGCTAAATTTGGCTCCATCTTTTCTCAGACGAATG 601
 QY 733 AATTGCAATCGAACCCCAAGTGAAAAAATTCCTGTAATTCGATGATGAGCT 792
 DB 600 AATTGCAATCGAACCCCAAGTGAAAAAATTCCTGTAATTCGATGATGAGCT 541
 QY 793 GCAGTACACAGATCTTACCGCTCCACAAAGTCAGAGATGTAATGCTCAATAGTG 852
 DB 540 GCAGTACACAGATCTTACCGCTCCACAAAGTCAGAGATGTAATGCTCAATAGTG 481
 QY 853 ACTTTTATTTATTCCTTGAATCAGACAGTAACTATTTTCAAGATGTTTAAAC 912
 DB 480 ACTTTTATTTATTCCTTGAATCAGACAGTAACTATTTTCAAGATGTTTAAAC 421
 QY 913 CTTTGTGCTGCTGTTATTAATATGCTGTAATCTTGTGCTTCTGATACAGAC 972
 DB 420 CTTTGTGCTGCTGTTATTAATATGCTGTAATCTTGTGCTTCTGATACAGAC 361
 QY 973 TGTTCCTGCTGCTGTTATTAATATGCTGTAATCTTGTGCTTCTGATACAGAC 1032
 DB 360 TGTTCCTGCTGCTGTTATTAATATGCTGTAATCTTGTGCTTCTGATACAGAC 301
 QY 1033 GTGAGGTTAGCTGCTGTAAGTGAAGTCAAGCTATTTGATCAACAGCAGACAGT 1092
 DB 300 GTGAGGTTAGCTGCTGTAAGTGAAGTCAAGCTATTTGATCAACAGCAGACAGT 241
 QY 1093 GTCTGCTCTTCCATGATGATTAAGTTTGTGAGATGTTATGATGAGATGCTGCTA 1152
 DB 240 GTCTGCTCTTCCATGATGATTAAGTTTGTGAGATGTTATGATGAGATGCTGCTA 181
 QY 1153 GTTCTGCTCTTGAAGTGAATTAAGTGAAGATTTACATCTGATTAAGTCTTCTAT 1212
 DB 180 GTTCTGCTCTTGAAGTGAATTAAGTGAAGATTTACATCTGATTAAGTCTTCTAT 121
 QY 1213 ACTGTCATATTAATTTGTGCTGAGAAATTTGTAATTTGTGACACTATGTAACAAA 1272
 DB 120 ACTGTCATATTAATTTGTGCTGAGAAATTTGTAATTTGTGACACTATGTAACAAA 61
 QY 1273 CAACGAGATATGTTAATAATTTGACTATTTGAGAGTAAAAAATTTAAAAA 1332
 DB 60 CAACGAGATATGTTAATAATTTGACTATTTGAGAGTAAAAAATTTAAAAA 1

 RESULT 3
 AA746643 811 bp mRNA linear EST 22-JAN-1998
 LOCUS nx27c08.s1 NCI CGAP GC4 Homo sapiens cDNA clone IMAGE:1257326 3'
 DEFINITION similar to gb:M28209 BAS-RELATED PROTEIN RAB-1A (HUMAN);, mRNA
 sequence.
 ACCESSION AA746643
 VERSION AA746643.1 GI:2786629
 KEYWORDS
 SOURCE human,
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap/
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.lnl.gov/db/rdp/image/image.html
 Insert length: 861 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 471.
 Location/Qualifiers
 1..811
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="1257326"
 /clone_lib="NCI-CGAP-GC4"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pRTT3D-Pac (Pharmacia) with a modified
 polylinker: 1st strand cDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pRTT3
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT 269 a 165 c 145 g 232 t
 ORIGIN
 Query Match 51.8%; Score 727.6; DB 9; Length 811;
 Best Local Similarity 95.8%; Pred. No. 9.8e-77;
 Matches 780; Conservative 0; Mismatches 29; Indels 5; Gaps 3;

QY 502 AGGAATTCCTGATTCCTTGGAATTCGTTTGGAAACCAAGCTGTAAGACACGA 561
 DB 809 AGGAATTCCTGATTCCTTGGAATTCGTTTGGAAACCAAGCTGTAAGACACGA 752
 QY 562 ATGTAAGACAGTCTTATGATGAGTGAATTAAGTGAATTAAGTGAATTAAGTGA 621
 DB 751 ATGTAAGACAGTCTTATGATGAGTGAATTAAGTGAATTAAGTGAATTAAGTGA 695
 QY 622 CAACAGCTGTGCTGCTGAGAGTCCATGTTAAATTTAGAGCACTCAGTCAAGCACT 681
 DB 694 CAACAGCTGTGCTGCTGAGAGTCCATGTTAAATTTAGAGCACTCAGTCAAGCACT 635
 QY 682 CAGGTGAGGTGTGCTGTAATTAATTTGCTCCATCTTTCTCAGCAATGAATTTGCA 741
 DB 634 CAGGTGAGGTGTGCTGTAATTAATTTGCTCCATCTTTCTCAGCAATGAATTTGCA 575
 QY 742 TCTGAACCAAGTGAATTAATTTGCTGTAATTTGCTGTAATTTGCTGTAATTTGCT 801
 DB 574 TCTGAACCAAGTGAATTAATTTGCTGTAATTTGCTGTAATTTGCTGTAATTTGCT 515
 QY 802 CAGATTCCTAGCGTCTCCAGAAAGTGAATTTGTAATTTGTAATTTGTAATTTGTA 861
 DB 514 CAGATTCCTAGCGTCTCCAGAAAGTGAATTTGTAATTTGTAATTTGTAATTTGTA 455
 QY 862 TTAATCCCTGACCTCAAGACAGTAACTTCAAGTGAATTTGTAATTTGTAATTTGTA 921
 DB 454 TTAATCCCTGACCTCAAGACAGTAACTTCAAGTGAATTTGTAATTTGTAATTTGTA 395
 QY 922 CTGTTTAAATTAATTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 981
 DB 394 CTGTTTAAATTAATTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTT 335
 QY 982 TGGTGGTGAATTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTA 1041
 DB 334 TGGTGGTGAATTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTA 275
 QY 1042 AGTCTTCTAGATGAATTCAGCAATTTGTAATTCACAGCAAGCAAGTGTCTGCAC 1101
 DB 274 AGTCTTCTAGATGAATTCAGCAATTTGTAATTCACAGCAAGCAAGTGTCTGCAC 215
 QY 1102 TTTCATGATTAATTTAGTGAATTTAGTGAATTTAGTGAATTTAGTGAATTTAGTGC 1161
 DB 214 TTTCATGATTAATTTAGTGAATTTAGTGAATTTAGTGAATTTAGTGAATTTAGTGC 155

QY 1162 TGTAGAGTTAAATGGAAGATTAACACTATCTGATTATAGTTCCTTCATACCTGCAAT 1221
|||||
Db 154 TGTAGAGTTAAATGGAAGATTAACACTATCTGATTATAGTTCCTTCATACCTGCAAT 95
QY 1222 ATAAATTTGGCTGCAGAAATATGTAAATTTGTTCGCACCTATGTACAAACACTGAAG 1281
|||||
Db 94 ATAAATTTGGCTGCAGAAATATGTAAATTTGTTCGCACCTATGTACAAACACTGAAG 35
QY 1282 ATATGTTAAATAATATGTACTTATTGGAAGTA 1315
|||||
Db 34 ATATGTTAAATAATATGTACTTATTGGAAGTA 1
RESULT 4
B1869965
LOCUS 817 bp mRNA linear EST 11-OCT-2001
DEFINITION 603394116F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5404275 5',
mRNA sequence.
ACCESSION B1869965
VERSION B1869965.1 GI:16043638
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Inocyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12030 row: P column: 04
High quality sequence stop: 669.
FEATURES
source
Location/Qualifiers
1..817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5404275"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-Sport6; site_1: NotI;
site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 237 a 155 c 169 g 256 t
ORIGIN
Query Match 51.5%; Score 723.8; DB 13; Length 817;
Best Local Similarity 95.3%; Pred. No. 2.7e-76;
Matches 779; Conservative 0; Mismatches 32; Indels 6; Gaps 3;
QY 475 AGAAGTAGTAGACTACACACAGCAGGAAGATTTGCTGATTCCTTGGAAATTCGTTT 534
|||||
Db 1 AGAAGTAGTAGACTACACACAGCAGGAAGATTTGCTGATTCCTTGGAAATTCGTTT 60
QY 535 TGGAAACAGAGCTAGAGTGAACGAATGTAGAACTCTTTCATGACATGGCAGCTG 594
|||||
Db 61 TGGAAACAGAGCTAGAGTGAACGAATGTAGAACTCTTTCATGACATGGCAGCTG 120
QY 595 AGATTAAAGCAGATGGGTCCGAGACACAGCTGTGTGAGAGTCCAAATGTTA 654
|||||
Db 121 AGATTAAAGCAGATGGGTCCGAGACACAGCTGTGTGAGAGTCCAAATGTTA 180

QY 655 AAATTCAGAGCAGCTCCAGCTCAAGCAGTCAGAGTTGCTGCTAAATTTGCCCTCAT 714
|||||
Db 181 AAATTCAGAGCAGCTCCAGCTCAAGCAGTCAGAGTTGCTGCTAAATTTGCCCTCAT 240
QY 715 CCTTTTCACAGCAATGCAATTTTGCATCTCAACCCCAAGTAAAAACCAATTTGCCCTGA 774
|||||
Db 241 CCTTTTCACAGCAATGCAATTTTGCATCTCAACCCCAAGTAAAAACCAATTTGCCCTGA 300
QY 775 ATTTGACTGTATGTAGCTGACACTCAACAGATTTCTTACCCGTCACAAAGTCAGAGAT 834
|||||
Db 301 ATTTGACTGTATGTAGCTGACACTCAACAGATTTCTTACCCGTCACAAAGTCAGAGAT 360
QY 835 TGTAAATGTCATATGACTGCTTTTATTTTCCCTGACTCAAGACGCTAATCTTAT 894
|||||
Db 361 TGTAAATGTCATATGACTGCTTTTATTTTCCCTGACTCAAGACGCTAATCTTAT 420
QY 895 TTCAGAACTGTTTAAACCTTTGTGTGCTGTTTATTAATAATGTGTATCTTGT 954
|||||
Db 421 TTCAGAACTGTTTAAACCTTTGTGTGCTGTTTATTAATAATGTGTATCTTGT 480
QY 955 GCTTTCCGTATACAGAGCTGTTCCCGGTGGTGGTGAATATATTTGTTGATGTT 1014
|||||
Db 481 GCTTTCCGTATACAGAGCTGTTCCCGGTGGTGGTGAATATATTTGTTGATGTT 540
QY 1015 ATATGGCATGTTAGATGTCAGGTTTATGCTCTGGAAGATGAAGTTCAGCCATTTGTA 1074
|||||
Db 541 ATATGGCATGTTAGATGTCAGGTTTATGCTCTGGAAGATGAAGTTCAGCCATTTGTA 600
QY 1075 TCAACAGACACAGCAGTGTCTGACCTTCCATGCAATGAAGTTAGTGAGATGTTAT 1134
|||||
Db 601 TCAACAGACACAGCAGTGTCTGACCTTCCATGCAATGAAGTTAGTGAGATGTTAT 660
QY 1135 GTAGATCT---GATTTGCTGATCTTCCCTGTGAGATTAAATGGAATTAACATA 1191
|||||
Db 661 GTAGATCTGATCTTCCCTGTGAGATTAAATGGAATTAACATA 720
QY 1192 TGTGATTAAATGTTCTTCACTCTGCAATATATTTGGTGCAGAA--TATGTAA 1248
|||||
Db 721 TGTGATTAAATGTTCTTCACTCTGCAATATATTTGGTGCAGAACTTCTGCAA 780
QY 1249 TTTGTGACACTATGTATCAAAACCACTGAAGTAT 1285
|||||
Db 781 TATGTGCCACACTATGTATCAAAACCACTGAAGTAT 817
RESULT 5
B0276678
LOCUS 910 bp mRNA linear EST 07-MAY-2002
DEFINITION AGENCOURT_7025320 NIH_MGC_126 Homo sapiens cDNA clone IMAGE:5809629
5', mRNA sequence.
ACCESSION B0276678
VERSION B0276678.1 GI:20486886
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2055 row: 1 column: 22
High quality sequence stop: 558.
FEATURES
Location/Qualifiers

	source	1. .910
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5809629"	
	/clone_1lb="NIH.MGC.126"	
	/tissue_type="mixed (pool of 40 RNAs)"	
	/lab_host="PH10B (T1-phage-resistant)"	
	/note="vector: pMIR-11b; Site_1: SfilI (ggcgcattggc); Site_2: SfilI (ggcgcctggc); Double-stranded cDNA was prepared from pool of 40 cell line poly(A) RNAs (bladder 48, blood 33.4%, brain 5.6%, breast 12.5%, colon 4%, connective tissue 1.4%, eye 1%, intestine 2.6%, kidney 2.2%, liver 5.7%, lung 1%, NK-cell 5.2%, prostate 4.3%, salivary gland 1.3%, and skin 2.3%). 5' and 3' adaptors were used in cloning as follows:	
	5'-AAGCAGTGTGATCAACGACAGATGAGGCGCATTCACGCCGG-3' and 5'-ATTGTAGAGCCGACGAGCGGCAGCATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech creator SMART kit and size-selected to contain the 0.5-1 kb size fraction (other fractions present in NIH.MGC.127 and NIH.MGC.128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC library."	
BASE COUNT	243 a 194 c 180 g 292 t 1 others	
ORIGIN		
Query Match	50.4%: Score 708.8; DB:14; Length 910;	
Best local similarity	98.9%: Pctd. No. 1.4e-74;	
Matches 724; Conservative	0; Mismatches 7; Indels 1; Gaps 1	
Db	584 GATGGCAGCTGAGATTTAAAGGAGTGGTCCCGGACAAAGAGTGGTGGCTGAGAA 643	
	3 GATGGCAGCTGAGATTTAAAGGAGTGGTCCCGGACAAAGAGTGGTGGCTGAGAA 62	
Qy	644 GTCCAACTGTTAAATTCAGACACTCCAGTCAAGCACTAGTGGAGGTGCTCTAAA 703	
	63 GTCCAACTGTTAAATTCAGACACTCCAGTCAAGCACTAGTGGAGGTGCTCTAAA 122	
Db	704 TTGGCTCCATCCCTTTCTCAGCAGCATGAAATTTGCATGTCAACCAAGTGAAGAAA 763	
	123 TTGGCTCCATCCCTTTCTCAGCAGCATGAAATTTGCATGTCAACCAAGTGAAGAAA 182	
Qy	764 AAATGGCTGAATTTACTGATGTAGTGCATCAACAGATTTTACGCTCCACAA 823	
	183 AAATGGCTGAATTTACTGATGTAGTGCATCAACAGATTTTACGCTCCACAA 242	
Db	824 AGCTCAGAGATGTGTAATGTCATCTACTCTTTTATTCCTCGACTCAAGACG 883	
	243 AGCTCAGAGATGTGTAATGTCATCTACTCTTTTATTCCTCGACTCAAGACG 302	
Qy	884 CTAACTTCATTTTCAGACACTGTTTAAACCTTTGTGCGGTATTAATTAATGTGG 943	
	303 CTAACTTCATTTTCAGACACTGTTTAAACCTTTGTGCGGTATTAATTAATGTGG 362	
Db	944 TAACTCCTTTGCTTCTGTAACAGAGTGTTCCTCGGAGTGGTTAGAAATATTTG 1003	
	363 TAACTCCTTTGCTTCTGTAACAGAGTGTTCCTCGGAGTGGTTAGAAATATTTG 422	
Qy	1004 TTTCGAGCTTAAATGAGCATGTTTGAAGTTCAGAGTTTACCTTCTGAAATTAAGTCA 1063	
	423 TTTCGAGCTTAAATGAGCATGTTTGAAGTTCAGAGTTTACCTTCTGAAATTAAGTCA 482	
Db	1064 GGCATTTTGTATCAACAGCAAGCATGTCTGTCACTTTCATGATCAATTAAGTTAGT 1123	
	483 GGCATTTTGTATCAACAGCAAGCATGTCTGTCACTTTCATGATCAATTAAGTTAGT 542	
Qy	1124 AGAGTTATATGTAAGATCTGATTTCTAGTTCTTCCTGTAGAGTTAATATGAAGAA 1183	
	543 AGAGTTATATGTAAGATCTGATTTCTAGTTCTTCCTGTAGAGTTAATATGAAGAA 602	
Db	1184 TTACAGTATCTGAATTAATAGTTCTTCACTCTGCATATAATTTGTGTGGTCAGAAAT 1243	

[illegible]

Db	754	TTTCATGACGAT - GCAGCGAGATTA	AAAAAGCG - ATTGGTCCGGAGCAACGCTGGT - G	698
OY	635	TGCTGAGAAGTC	CCATGTTTAAATTCAGAGCACTCCAGTCAAGCAAGTCAGGTGAGTTG	694
Db	697	TGCTGAGAAAGTC	CCATGTTTAAATTCAGAGCACTCCAGTCAAGCAAGTCAGGTGAGTTG	638
OY	695	CTGGCAAAATTTGGCTCATCCTTTGTCACAGCAATGAATTTGCAATCTGAACCAAGT		754
Db	637	CTGCTAAATTTGGCTCATCCTTTTCTCACAGCAATGAA - TTGCAATCTGAACCAAGT		579
OY	755	GAACAAACAAATTTGCTGAAATTTGACTGATAGTACGCACTACAAACAGATTTCTACCG		814
Db	578	GAACAAACAAATTTGCTGAAATTTGACTGATAGTACGCACTACAAACAGATTTCTACCG		519
OY	815	TTCTCCAAAGGTCAGACATTTGTAATGTCGAATTTCTGACTTTTATTCCTTGAC		874
Db	518	TTCTCCAAAGGTCAGACATTTGTAATGTCGAATTTCTGACTTTTATTCCTTGAC		459
OY	875	TCAACAGACGCTAACCTCAATTTTTCACAACGCTTTTAAACCTTTGTGCTGCTTATATAA		934
Db	458	TCAACAGACGCTAACCTCAATTTTTCACAACGCTTTTAAACCTTTGTGCTGCTTATATAA		399
OY	935	TAAATGTGTAACTCCTTGCTGCTTCCGTGATACCAAGACTGTTCCGCTGGTGGTTAGAA		994
Db	398	TAAATGTGTAACTCCTTGCTGCTTCCGTGATACCAAGACTGTTCCGCTGGTGGTTAGAA		339
OY	995	TATATTTTGTTTGATGTTTATATGAGCATGTTTATGATGTCAGGTTTATGCTTCGAGAA		1054
Db	338	TATATTTTGTTTGATGTTTATATGAGCATGTTTATGATGTCAGGTTTATGCTTCGAGAA		279
OY	1055	TGAAGTTCAGCCATTTTGTATCAAAACAGCACAGCAAGCTGTCAGTCTTCATGSCATPA		1114
Db	278	TGAAGTTCAGCCATTTTGTATCAAAACAGCACAGCAAGCTGTCAGTCTTCATGSCATPA		219
OY	1115	AGTTTAGAGAGATGTTATATGTAATGTAACATCTGATTTGCTAGTTCTTCCCTTGACATTATA		1174
Db	218	AGTTTAGAGAGATGTTATATGTAATGTAACATCTGATTTGCTAGTTCTTCCCTTGACATTATA		159
OY	1175	ATGCAAAATTAACATATCTGATTTGAATAGTTTCTTCATCTCTGCAATATAATTTGGGCT		1234
Db	158	ATGCAAAATTAACATATCTGATTTGAATAGTTTCTTCATCTCTGCAATATAATTTGGGCT		99
OY	1235	GCAGAAATATTTGTAATTTGTCACACTATGTAACAAACAACTGGAAGATATGTTAATPA		1294
Db	98	GCAGAAATATTTGTAATTTGTCACACTATGTAACAAACAACTGGAAGATATGTTAATPA		39
OY	1295	ATATGTACTTATTTGAGAGTAAAAA	AAAAAAAAAAAA 1332	
Db	38	ATATGTACTTATTTGAGAGTAAAAA	AAAAAAAAAAAA 1	

```

RESULT 7
AL530265
LOCUS
DEFINITION AL530265 LTI_NF1001_NBC4 Homo sapiens cDNA clone CSDD009YH19 5
ACCESSION AL530265
VERSION AL530265
KEYWORDS AL530265.1 GI:12793758
SOURCE EST.
ORGANISM human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope -- Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segr@genoscope.cns.fr, Web : www.genoscope.cns.fr.
location/Qualifiers
FEATURES

```

```

source
1. .945
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GSDDD009YH19"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: liang@lifetech.com URL: http://fulllength.invitrogen.com"
BASE COUNT 280 a 186 c 224 g 246 t 9 others
ORIGIN

```

Query Match	50.4%;	Score 707.8;	DB 9;	Length 945;
Best Local Similarity	88.4%;	Pred. No. 1.8e-74;		
Matches 835;	Conservative 5;	Mismatches 7;	Indels 98;	Gaps 3;
QY 60	CGTTTGCCTCCCGGAGACGCCCTATCTCATCTTCTTTCGATTACCCGTGGCGCGAG	119		
Db 1	CGTTTGCCTCCCGGAGACGCCCTATCTCATCTTCTTTCGATTACCCGTGGCGCGAG	60		
QY 120	AGTCAGGGCGCGGCTGGCGCAGCAAGGGCGGGGTGGCGCGGGGAGCTGCAGTAC	179		
Db 61	AGTCAGGGCGCGGCTGGCGCAGCAAGGGCGGGGTGGCGCGGGGAGCTGCAGTAC	120		
QY 180	ATGTCCACATGAATCCCGAATATGATTTATTTCAAGTTACTCTGATTTGGCAGTCA	239		
Db 121	ATGTCCACATGAATATCCCGAATATGATTTATTTCAAGTTACTCTGATTTGGCAGTCA	180		
QY 240	GGGGTTGGAAGATCTTGGCTTCTCTTAGGTTTGAGATGATACATATACAGAAAGCTAC	299		
Db 181	GGGGTTGGAAGATCTTGGCTTCTCTTAGGTTTGAGATGATACATATACAGAAAGCTAC	240		
QY 300	ATCAGCACAAATTTGGTGTGGATTTCAAATTAAGACTATAGAGTTAGACGGGAAAAATC	359		
Db 241	ATCAGCACAAATTTGGTGTGGATTTCAAATTAAGACTATAGAGTTAGACGGGAAAAATC	300		
QY 360	AAGCTTTCAATA-----	371		
Db 301	AAGCTTCAAAATTTGGGACACACAGCGCCAGCAAGATTTCCGACATATCACTCCAGTTAT	360		
QY 372	-----GAGTCCCTTCAT	383		
Db 361	TACAGAGAGGCCCATGGCATCAATGTTGTGATGATGATGACAGATCAGAGAGTCTTCAT	420		
QY 384	AATGTTAAACAGTGGCTGCAGGAATAGATGCTTATGCGCAGTGAATAATGTCACAAATTG	443		
Db 421	AATGTTAAACAGTGGCTGCAGGAATAGATGCTTATGCGCAGTGAATAATGTCACAAATTG	480		
QY 444	TTGGTAGGGAACAAATGATCTGTGACCAACAAGAAAGTAGACTATACACAACAGCGAG	503		
Db 481	TTGGTAGGGAACAAATGATCTGTGACCAACAAGAAAGTAGACTATACACAACAGCGAG	540		
QY 504	GAATTTGCTGATTCCTCTTGGAAATTCGTTTTTGGAAACACAGTGTAAAGATGCAACGAT	563		
Db 541	GAATTTGCTGATTCCTCTTGGAAATTCGTTTTTGGAAACACAGTGTAAAGATGCAACGAT	600		
QY 564	GTAGAACAGTCTTTCATGACGATGCGACGTGAGATTAATAAGCGAATGGTCCGGAGCA	623		
Db 601	GTAGAACAGTCTTTCATGACGATGCGACGTGAGATTAATAAGCGAATGGTCCGGAGCA	659		
QY 624	ACAGCTGGTGGTGTGAGAAATGTCAAATGTTAAATTCAGACACATCCAGTCAAGCAGTCA	683		
Db 660	ACAGCTGGTGGTGTGAGAAATGTCAAATGTTAAATTCAGACACATCCAGTCAAGCAGTCA	719		

```

Oy 684 -GTGAGAGTGTGCTGCTAAATTTGGCCATCCCTTTTCTCAGACGAATGATTTGCAAT 742
      |||
Db 720 GGTGTGAGTGTGCTGCTAAATTTGGCCATCCCTTTTCTCAGACGAATGATTTGCAAT 779
      |||
Oy 743 CTGAACCCAGTGAATAAAATTTGCCGTAATTTGCTATGATAGCTGCATACAC 802
      |||
Db 780 CTGAACCCAGTGAATAAAATTTGCCGTAATTTGCTATGATAGCTGCATACAC 839
      |||
Oy 803 AGATCTTACCGTCTCCACAAAGCTCAGAGATTTGAATGCTCAATAGCTATTTT 862
      |||
Db 840 AGATCTTACCGTCTCCACAAAGCTCAGAGATTTGAATGCTCAATAGCTATTTT 899
      |||
Oy 863 TATTCCTGACTCAGACAGCTACTTCTATTTTCAGAACTGTT 907
      |||
Db 900 TATTCCTGACTCAGACAGCTACTTCTATTTTCAGAACTGTT 944
      |||

RESULT 8
LOCUS BM452262 1105 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6386191 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5526635
ACCESSION BM452262
VERSION BM452262.1 GI:18501302
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1105)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bhs-remail.nih.gov
Tissue Procurement: ATCC/DC/D/DFP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM12200 row: 1 column: 12
High quality sequence stop: 733.
Location/Qualifiers
1. 1105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5526635"
/clone_1id="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/label="note:DH10B (phage-resistant)"
/site_2="Skin: Vector: PCMV-SPOF6; site_1: NotI;
site_2: SalI; Cloned unidirectionally; primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 320 a 216 c 279 g 282 t 8 others
ORIGIN
Query Match 50.1%; Score 704.4; DB 13; Length 1105;
Best Local Similarity 86.0%; Pred. No. 4.1e-74;
Matches 899; Conservative 0; Mismatches 43; Indels 103; Gaps 7;

Oy 1 AAGCGATAGCTGACT- GCGGCGCTGCTGATGTTGTTCTAGAGGACGAGTAGGGGAAGA 59
      |||
Db 2 AAGCGATAGCTGAGTGGCGGCTGCTGATGTTGTTCTAGAGGACGAGTAGGGGAAGA 61
      |||
Oy 60 CGTTTGTCTCCCGGACAGCCCTATCTCATCTCTTTTGATATACCCGTGGCGGAG 119
      |||
Db 62 CGTTTGTCTCCCGGACAGCCCTATCTCATCTCTTTTGATATACCCGTGGCGGAG 121
      |||
Oy 120 AGTCAGGCGCGGCTGCGGACGACAGGCGGCTGCGCGGCGGACGATGACAGTGCAC 179
      |||

Db 122 AGTCAGGCGCGGCTGCGGACGACAGGCGGCGGCTGCGCGGCGGACGATGACAGTGCAC 181
      |||
Oy 180 ATGTCAGATGAATTCGGAATATGATTTATTTATTCAGTACTTGTGATGGCAGCTCA 239
      |||
Db 182 ATGTCAGATGAATTCGGAATATGATTTATTTATTCAGTACTTGTGATGGCAGCTCA 241
      |||
Oy 240 GGGGTGGAAGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 299
      |||
Db 242 GGGGTGGAAGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 301
      |||
Oy 300 ATGACACAAATTTGGTGGATTTCAAAATGAAGACTATAGCTTGAAGGGAAGAAATAC 359
      |||
Db 302 ATGACACAAATTTGGTGGATTTCAAAATGAAGACTATAGCTTGAAGGGAAGAAATAC 361
      |||
Oy 360 AAGCTTCAAAATTTGGTGGATTTCAAAATGAAGACTATAGCTTGAAGGGAAGAAATAC 371
      |||
Db 362 AAGCTTCAAAATTTGGTGGATTTCAAAATGAAGACTATAGCTTGAAGGGAAGAAATAC 421
      |||
Oy 372 -----GAGTCTTCAT 383
      |||
Db 422 TACAGACGACCCCATGGCATCATAGTTGTATGATGTACAGATCAGAGAGTCTTCAT 481
      |||
Oy 384 AATGTTAAACAGTGGCTGACGGAATATGATGCTTATGCCAGTGAATGCAACAAATTTG 443
      |||
Db 482 AATGTTAAACAGTGGCTGACGGAATATGATGCTTATGCCAGTGAATGCAACAAATTTG 541
      |||
Oy 444 TTGTTAGGGAACAAATGTGATCTGACCAAAAGAAAGTAGTAGCTACACAGGGAAG 503
      |||
Db 542 TTGTTAGGGAACAAATGTGATCTGACCAAAAGAAAGTAGTAGCTACACAGGGAAG 601
      |||
Oy 504 GAATTTGCTGATTCCTTTGGAAATTCGTTTGGAAACGACGCTAAGAAATGCAACGAAT 563
      |||
Db 602 GAATTTGCTGATTCCTTTGGAAATTCGTTTGGAAACGACGCTAAGAAATGCAACGAAT 661
      |||
Oy 564 GTGAACAGTCTTTTCATGACAGATGACAGCTGATTTAAACGCAATGGTCCCGGAGCA 623
      |||
Db 662 GTGAACAGTCTTTTCATGACAGATGACAGCTGATTTAAACGCAATGGTCCCGGAGCA 721
      |||
Oy 624 ACAGCTGTGTGCTGAGAAAGTCAATGTTAAATTCAGAGCAGTCCAGTCAAGCAGTCA 683
      |||
Db 722 ACAGCTGTGTGCTGAGAA-TCGAATGTTAAATTCAGAGCAGTCCAGTCAAGCAGTCA 780
      |||
Oy 684 -GGTGAGGTTGCTGCTAAATTTGCTCCTCATCCCTTTTCTCAGACGAATGATTTGCAAT 742
      |||
Db 781 NGGTGAGGTTGCTGCTAAATTTGCTCCTCATCCCTTTTCTCAGACGAATGATTTGCAAT 840
      |||
Oy 743 CTGAACCCAGTGAATAAAATTTGCCGTAATTTGCTATGATGATGCTGCATCAAC 802
      |||
Db 841 TGAACCCAGTGAATAAAATTTGCCGTAATTTGCTATGATGATGCTGCATCAAC 899
      |||
Oy 803 AGATTTTACCGTCTCCACAAAGCTCAGAGATTTGAATGCTCAATAGCTG- ACTTTT 860
      |||
Db 900 AGATTTTACCGTCTCCACAAAGCTCAGAGATTTGAATGCTCAATAGCTGACTTTT 959
      |||
Oy 861 TTTATTCCTGACTCAGACAGCTACTTCTATTTTCAGACTGTTTAAAC- TTGTG 919
      |||
Db 960 TTTATTCCTGACTCAGACAGCTACTTCTATTTTCAGACTGTTTAAACCTTTTGTG 1019
      |||
Oy 920 TGCTGTTTAAATATGCTGT 944
      |||
Db 1020 TGCTGTTTAAATATGCTGT 1044
      |||

RESULT 9
LOCUS BQ448090/c 725 bp mRNA linear EST 29-MAY-2002
DEFINITION UI-H-BD1-bal-k-14-0-UI.s1 NCI_CGAP_Ccl1 Homo sapiens cDNA clone
ACCESSION BQ448090
VERSION BQ448090.1 GI:21251202
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
```



```

BASE COUNT      235 a      140 c      168 g      253 t
ORIGIN           this is a NIH_MGC Library.

```

Query Match	50.0%;	Score 702;	DB 13;	Length 796;
Best Local Similarity	95.7%;	Pred. No. 9.8e-74;		
Matches 754; Conservative	0;	Mismatches 30;	Indels 4;	Gaps 3

[illegible]

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	1 (bases 1 to 760)
TITLE	Bonaldi,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
COMMENT	Contact: McCray, PB McCray Lab University of Iowa 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA Tel: 319 356 4866 Fax: 319 356 7171 Email: paul-mccray@uiowa.edu Tissue Procurement: Dr. M. J. Welsh, university of Iowa CDNA library preparation: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa DNA Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) . Seq primer: M13 FORWARD POLYA=yes.
FEATURES	Location/Qualifiers 1..760 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="UI-CF-EC1-ab1-a-17-0-01" /clone_11b="UI-CF-EC1" /clone_11b="UI-CF-EC1" /tissue_type="lung" /dev_stage="Adult and Petal" /lab_host="DH10B (Life Technologies) (T1 phage resistant)" /note="Organ: Lung, Vector: pRTR-Pac (Pharmacia) with a modified polylinker, Site-1, Ecor I, Site-2, Not I; UI-CF-EC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pRTR-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d7)18 tail. The sequence tag for this library is AAGGCTTAC. TAG_L1B=UI-CF-EC1 TAG_T1S2=Normal lung Epithelial Cells Tissue nos 369-371 and 380-383 TAG_SEQ=AAGGCTTAC"
BASE COUNT	255 a 142 c 125 g 236 t 2 others
ORIGIN	
Query Match	49 38; Score 692.8; DB 14; Length 760;
Best Local Similarity	99 18; Pseq No 1,2e-72;
Matches 738:	Conservative 0; Mismatches 3; Indels 4; Gaps 4;
Y	588 GCACCTAGATTAAAAAGCAATGSGTCCCGAGCAACAGCTGGTGTGTGACGAAGTCC 647
Db	741 GACGCTAGAGTTAAAAAGCAAT-6GTCGCGAGCAACAGCT-6TGCGMCTGAGAAGTCC 664
Y	648 AATGTTAAATTCAGAGCACTCCAGTCAGAGCTGAGTGGAGGTGCTGCTAAAAATTG 707
Db	683 AATGTTAAATTCAGAGCACTCCAGTCAGAGCTGAGTGGAGG-TGCGTCTTAAA-TTG 626
Y	708 CCTCATCTCTTTTCACAGCAATGAATTTGCATCTGCAACCCAAAGTGAACAAAT 767
Db	625 CCTCATCTCTTTTCACAGCAATGAATTTGCATCTGCAACCCAAAGTGAACAAAT 566
Y	768 TGCGTGAATTTGACTGATATGACTGCACTCAACAGATCTTCCGCTCCACAAAGGT 827

```

|||||
Db 565 TGCGTGAATGACTGATGAGTGCACACACAGATCTTACCGCTCCACAAAGT 506
QY 828 CAGAGATGTAATGATCACTGACTTTTTTTTATTCCTGACGCAAGACGCTAA 887
Db 505 CAGAGATGTAATGATCACTGACTTTTTTTTATTCCTGACGCAAGACGCTAA 446
QY 888 CTTCATTTTGAAGACTGTTTAAACCTTGTGCTGCTGTTTAAATTAATGCTGTAA 947
Db 445 CTTCATTTTGAAGACTGTTTAAACCTTGTGCTGCTGTTTAAATTAATGCTGTAA 386
QY 948 CCTGTGCTGCTTCTGATACAGACTGTTCCCTGCTGCTGTTGAATATATTTGTTT 1007
Db 385 CCTGTGCTGCTTCTGATACAGACTGTTCCCTGCTGCTGTTGAATATATTTGTTT 326
QY 1008 GATGTTATATGTCATGTTAGATGTCAGGTTAGTCTTCTGAGATGGAAGTTCAGCA 1067
Db 325 GATGTTATATGTCATGTTAGATGTCAGGTTAGTCTTCTGAGATGGAAGTTCAGCA 266
QY 1068 TTTTGTATCAAAACGACACAGCAGTGTCTGCTACCTTCCATGCAATAAAGTTAGTAGAT 1127
Db 265 TTTTGTATCAAAACGACACAGCAGTGTCTGCTACCTTCCATGCAATAAAGTTAGTAGAT 206
QY 1128 GTTATATGTAAGATCTGATTTGCTAGTCTTCTGCTGCTGTTAAATGGAAGATTAC 1187
Db 205 GTTATATGTAAGATCTGATTTGCTAGTCTTCTGCTGCTGTTAAATGGAAGATTAC 146
QY 1188 ACATGCTGATTAATAGTTTCTTCTCATCTGCTCATATATATTTGGCTGCGAGATTTGTA 1247
Db 145 ACATGCTGATTAATAGTTTCTTCTCATCTGCTCATATATATTTGGCTGCGAGATTTGTA 86
QY 1248 ATTTGTGTCACACTATGTACAAACACAGTGAATATGTTTAAATATATTTACTTAT 1307
Db 85 ATTTGTGTCACACTATGTACAAACACAGTGAATATGTTTAAATATATTTACTTAT 26
QY 1308 TGGAACTAAAAAATAAAAA 1332
Db 25 TGGAACTAAAAAATAAAAA 1

RESULT 12
BM919860 976 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6708189 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5749635
DEFINITION 5', mRNA sequence.
ACCESSION BM919860
VERSION BM919860.1 GI:19370239
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgarbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LHAM12779 row: f column: 04
High quality sequence stop: 757.
Location/Qualifiers
1. 976
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5749635"
/clone_lib="NIH_MGC_120"

```

```

/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH-MGC library."
BASE COUNT 285 a 196 c 246 g 248 t 1 others
ORIGIN
Query Match 49.2%; Score 691.8; DB 14; Length 976;
Best Local Similarity 87.6%; Pred. No. 1.3e-72;
Matches 832; Conservative 0; Mismatches 18; Indels 100; Gaps 3;

```

```

QY 16 GCGCGGCTGCTGATTTGTCTTCTAGGGGAGGAGTACGGGAACGTTGCTCCCGGA 75
Db 13 GCGCGGCTGCTGATTTGTCTTCTAGGGGAGGAGTACGGGAACGTTGCTCCCGGA 72
QY 76 ACAGCTATCATCTCTTCTTCTGATTTACCGCTGGCGGAGATCAGGGGCGGCT 135
Db 73 ACAGCTATCATCTCTTCTTCTGATTTACCGCTGGCGGAGATCAGGGGCGGCT 132
QY 136 GCGGACAGAGGGGCGGCTGCGCGGCGGCGGAGCTGACATGATCCAGATGATC 195
Db 133 GCGGACAGAGGGGCGGCTGCGCGGCGGCGGAGCTGACATGATCCAGATGATC 192
QY 196 CCGAATATGATTTATTTATTTCAAGTTACTTCTGATGCGGCTCAGGGGTGGAAGTCTT 255
Db 193 CCGAATATGATTTATTTATTTCAAGTTACTTCTGATGCGGCTCAGGGGTGGAAGTCTT 252
QY 256 GCGTCTCTTAGGTTGTCAGATGATATATACAGAACTCATAGCAATGAGCAATGGT 315
Db 253 GCGTCTCTTAGGTTGTCAGATGATATATACAGAACTCATAGCAATGAGCAATGGT 312
QY 316 TGGATTTCAAAATTAAGACTATAGAGTTAGACGGGAACCAATCAACCTTCAATA 371
Db 313 TGGATTTCAAAATTAAGACTATAGAGTTAGACGGGAACCAATCAACCTTCAATA 372
QY 372 ----- 371
Db 373 ACACAGCAGGCGCAGAAAGATTTGCAACATCACCTTACTTATACAGAGGCCCATG 432
QY 372 -----GAGTCCCTCAATATGTTAAACAGTGGC 399
Db 433 GCATCATAGTTGTGTATGATGTGACAGATCAGAGTCTTCAATATGTTAAACAGTGGC 492
QY 400 TGCAGAAATAGATCGTTATGCGCAGTGAATATGTCACAAATTTGTTGTAAGGAACAA 459
Db 493 TGCAGAAATAGATCGTTATGCGCAGTGAATATGTCACAAATTTGTTGTAAGGAACAA 552
QY 460 GTGATCTGACCACAAAGATAGTACATACACACAGCGAAGAAATTTGCTGATCC 519
Db 553 GTGATCTGACCACAAAGATAGTACATACACACAGCGAAGAAATTTGCTGATCC 612
QY 520 TTGAATTCGCTTTTGGAAACAGTGTCTAAGATGCAAGAAATGTAAGACAGTCTTCA 579
Db 613 TTGAATTCGCTTTTGGAAACAGTGTCTAAGATGCAAGAAATGTAAGACAGTCTTCA 672
QY 580 TGAAGATGAGCAGTGAATTAAGAAAGCAATGGTCCCGGAGCAACAGCTGTTGCTG 639
Db 673 TGAAGATGAGCAGTGAATTAAGAAAGCAATGGTCCCGGAGCAACAGCTGTTGCTG 732
QY 640 AGAAGTCAATGTTAAATTCAGAGCACTCCAGTCAAGCACTCAGGTGAGTGTGCT 699
Db 733 AGAAGTCAATGTTAAATTCAGAGCACTCCAGTCAAGCACTCAGGTGAGTGTGCT 792
QY 700 AAAATTTGCGTCATCTTCTTCTCAGCAATGATTTGAATCTGAACCCCAAGTGA 759
Db 793 AAAATTTGCGTCATCTTCTTCTCAGCAATGATTTGAATCTGAACCCCAAGTGA 852

```

```
QY 760 AACAAATGCTGTAATGTAAGTACAGACAG-ATTCTTACCGTCTC 818
DB 853 AACAAATGCTGTAATGTAAGTACAGACAG-ATTCTTACCGTCTC 912
QY 819 CACAAAGTCAGAGA---TTGTAATGTCATACACTGCTTTTAT 865
DB 913 CACAAAGTCAGAGAATGTAAATGTCATACACTGCTTTTAT 962

RESULT 13
BIOS6445 837 bp mRNA linear EST 20-JUN-2001
LOCUS 602849703F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4991233 5',
DEFINITION mRNA sequence.
ACCESSION BIOS6445
VERSION BIOS6445.1 GI:14504775
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.ncl.nih.gov/
AUTHORS 1 (bases 1 to 837)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
TITRE Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11008 row: f column: 02
High quality sequence stop: 758.
Location/Qualifiers
1..837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4991233"
/clone_lib="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: PCMV-SPORE6; Site 1: Nott;
Site 2: Salt; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 248 a 134 c 174 g 281 t
ORIGIN
Query Match 49.28; Score 690.6; DB 13; Length 837;
Best Local Similarity 97.58; Pred No 2; 1e-72;
Matches 754; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

QY 557 AACGAATGTGAACAGCTTCATGACAGACAGCTGAGATTAAAAAGCAATGGCTCC 616
DB 1 AACGAATGTGAACAGCTTCATGACAGACAGCTGAGATTAAAAAGCAATGGCTCC 60
QY 617 CGAGCAACAGCTGCTGCTGGAAGTCAATGTAATTCAGACAGCTCCAGTCAA 676
DB 61 CGAGCAACAGCTGCTGCTGGAAGTCAATGTAATTCAGACAGCTCCAGTCAA 120
QY 677 GCAGTCAGGTGAGGTGCTGCTAAATTTGCTCCATCTTTTTCACAGCAATGAAT 736
DB 121 GCAGTCAGGTGAGGTGCTGCTAAATTTGCTCCATCTTTTTCACAGCAATGAAT 180
QY 737 TGCATCTGAACCAAGTGAATAAATTCCTGTAATGTATGATGAGCTGCAC 796
DB 181 TGCATCTGAACCAAGTGAATAAATTCCTGTAATGTATGATGAGCTGCAC 240
QY 797 TACAAAGATTTCTACCGTCTCCACAAAGCTGAGATTTGTAATGATCATCTGACTT 856

FEATURES
source
1..837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4991233"
/clone_lib="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: PCMV-SPORE6; Site 1: Nott;
Site 2: Salt; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 248 a 134 c 174 g 281 t
ORIGIN
Query Match 49.28; Score 690.6; DB 13; Length 837;
Best Local Similarity 97.58; Pred No 2; 1e-72;
Matches 754; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

DB 241 TACACAGATTCCTTACCCTCCACAAAGTCAGAGATTTGTAATGTCATACGACTT 300
QY 857 TTTTTCCTTCCCTTGACATCAGACAGCTAACTTCATTTTCAGAACTGTTTAAACCTT 916
DB 301 TTTTTCCTTCCCTTGACATCAGACAGCTAACTTCATTTTCAGAACTGTTTAAACCTT 360
QY 917 GTGTGCTGTTTAAATTAATGTCATACCTTGTGCTTCCCTGATACAGACTGTT 976
DB 361 GTGTGCTGTTTAAATTAATGTCATACCTTGTGCTTCCCTGATACAGACTGTT 420
QY 977 TCCCGTGTGTTGATGATATATTTTGTGATGTTTATTTATGTCATGTTAGATGTA 1036
DB 421 TCCCGTGTGTTGATGATATATTTTGTGATGTTTATTTATGTCATGTTAGATGTA 480
QY 1037 GCTTATGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1095
DB 481 GCTTATGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
QY 1096 TGTGATGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1155
DB 541 TGTGATGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 1156 CTTCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213
DB 601 CTTCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 1214 CTCTGATTAATTTGTGCTGCAAGATATTTGATTTGTCACACTGATGATGATG 1273
DB 661 CTCTGATTAATTTGTGCTGCAAGATATTTGATTTGTCACACTGATGATGATG 719
QY 1274 AACTGAGATATGTTTAAATTAATTTGATTTGATTTGATTTGATTTGATTTGAT 1326
DB 720 AA-TGAGATATGTTTAAATTAATTTGATTTGATTTGATTTGATTTGATTTGAT 771

RESULT 14
BIOS6159 689 bp mRNA linear EST 28-FEB-2002
LOCUS BM706159
DEFINITION UI-E-DW0-agg-f-10-0-UI.r1 UI-E-DW0 Homo sapiens cDNA clone
ACCESSION UI-E-DW0-agg-f-10-0-UI 5', mRNA sequence.
VERSION BM706159
KEYWORDS EST.
SOURCE 19019417
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Ronald M.F., Lennon G. and Soares M.B.
TITRE Normalization and subtraction: two approaches to facilitate gene
JOURNAL discovery
MEDLINE Gene Res. 6 (9), 791-806 (1996)
COMMENT 9704447
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8280
Fax: 319 335 9265
Email: msocares@iui.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seg primer: M13 Reverse.
Location/Qualifiers
1..689
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-DW0-agg-f-10-0-UI"
/clone_lib="UI-E-DW0"
```

```

/tissue_type="lens"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: eye; Vector: pRT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UT-E-DM0 is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pRT3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGATTACGGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

```

BASE COUNT      195 a      122 c      142 g      230 t
ORIGIN
```

```
Query Match      49.0%; Score 689; DB 14; Length 689;
```

```
Best Local Similarity 100.0%; Pred. No. 3.6e-72;
Matches 689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 503 GGAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTAGAAATGCAACGAA 562
    1 GGAATTTGCTGATTCCTTGGAAATTCGTTTGGAAACAGTGTAGAAATGCAACGAA 60
QY 563 TGTAGAACAGCTTTCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 622
    61 TGTAGAACAGCTTTCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 623 AACAGCTGTGTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 682
    121 AACAGCTGTGTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 683 AGGTGAGAGTTGCTGCTAAATTTGGCTCCATCTTTTCTCAGACAGCAATTTGCAAT 742
    181 AGGTGAGAGTTGCTGCTAAATTTGGCTCCATCTTTTCTCAGACAGCAATTTGCAAT 240
QY 743 CTGAACCCAAAGTGAAGAAACAAATTTGCCAAATTTGCTATGCTATGCTATGCTATG 802
    241 CTGAACCCAAAGTGAAGAAACAAATTTGCCAAATTTGCTATGCTATGCTATGCTATG 300
QY 803 AGATTTCTACCGTCTCCACAAAGTGTAGAGATTTGAATGCTCAATCTGACTTTTCTT 862
    301 AGATTTCTACCGTCTCCACAAAGTGTAGAGATTTGAATGCTCAATCTGACTTTTCTT 360
QY 863 TATTCCTTGAAGTCAAGACAGCTAATCTTCAAGACTGTTTAAACCTTTGTGTGC 922
    361 TATTCCTTGAAGTCAAGACAGCTAATCTTCAAGACTGTTTAAACCTTTGTGTGC 420
QY 923 TGGTTTAAATTAATGATGCTATCTTGTGCTTCCATATACAGACTGTTTCCGT 982
    421 TGGTTTAAATTAATGATGCTATCTTGTGCTTCCATATACAGACTGTTTCCGT 480
QY 983 GGTGTTAGAAATATATTTTGTGATGTTATTTGAGTGTGTTAGATGTCAAGTTTA 1042
    481 GGTGTTAGAAATATATTTTGTGATGTTATTTGAGTGTGTTAGATGTCAAGTTTA 540
QY 1043 GTCCTTGAAGATGAAGTTCAGCATTCTTATCAAAACAGACAGAGAGTCTGTCACT 1102
    541 GTCCTTGAAGATGAAGTTCAGCATTCTTATCAAAACAGACAGAGAGTCTGTCACT 600
QY 1103 TTTCATGATTAAGTATTTAGTGAATGTTATTTAGTGAATGCTGATTTGCTTCTCTT 1162
    601 TTTCATGATTAAGTATTTAGTGAATGTTATTTAGTGAATGCTGATTTGCTTCTCTT 660
QY 1163 GTAGAGTTAATTAAGTGAAGATTTACACTA 1191
    661 GTAGAGTTAATTAAGTGAAGATTTACACTA 689
```

```

RESULT 15
B1092003      737 bp      mRNA      linear      EST 20-JUN-2001
LOCUS      602856302.F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4997688 5',
DEFINITION      mRNA sequence.
ACCESSION      B1092003
VERSION      B1092003.1 GI:14510333
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 737)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgads-femail.nih.gov
            Tissue Procurement: ARCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: ILAM11025 row: c column: 01
            High quality sequence stop: 727.
```

```

FEATURES
    source
        1..737
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4997688"
            /clone_1id="NIH_MGC_10"
            /cell_line="MGC36"
            /lab_host="DH10B"
            /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.5 kb. Library prepared by Life
            Technologies."
```

```

BASE COUNT      214 a      135 c      147 g      241 t
ORIGIN
```

```
Query Match      49.0%; Score 688.2; DB 13; Length 737;
```

```
Best Local Similarity 97.8%; Pred. No. 4.3e-72;
Matches 719; Conservative 0; Mismatches 13; Indels 3; Gaps 2;
```

```

QY 550 AGATGCAACGAATGTAGAACAGCTTTTCATGACGATGGCAGCTGACATTTAAAGGAA 609
    1 AGATGCAACGAATGTAGAACAGCTTTTCATGACGATGGCAGCTGACATTTAAAGGAA 60
QY 610 TGGTCCCGAGCAACAGCTGTGCTGCTGAGAGTCCATGTAAATTCAGAGCACTC 669
    61 TGGTCCCGAGCAACAGCTGTGCTGCTGAGAGTCCATGTAAATTCAGAGCACTC 120
QY 670 CAGTCAAGCACTCAGGTGAGAGTGTCTGCTAAATTTGCCCTCAATCTTTCTCAGACA 729
    121 CAGTCAAGCACTCAGGTGAGAGTGTCTGCTGCTAAATTTGCCCTCAATCTTTCTCAGACA 180
QY 730 ATGAATTTGCAATGTGAACCAAGTGAAGAAACAAATTTGCTCAATTTGACTATGTA 789
    181 ATGAATTTGCAATGTGAACCAAGTGAAGAAACAAATTTGCTCAATTTGACTATGTA 240
QY 790 GCTGCACATCAAGATTTCTTACCGTCTCCACAAAGTGTAGAGATTTGAATGCTCATATA 849
    241 GCTGCACATCAAGATTTCTTACCGTCTCCACAAAGTGTAGAGATTTGAATGCTCATATA 300
QY 850 CTGACTTTTATTTATTCCTTGTGACTCAAGACAGCTAATCTTTCAGAACTGTTTAA 909
    301 CTGACTTTTATTTATTCCTTGTGACTCAAGACAGCTAATCTTTCAGAACTGTTTAA 360
QY 910 AACCTTTGTGTGCTGTTTAAATTAATGTGTGTAATCTTGTGCTTCTGATACCA 969
    AACCTTTGTGTGCTGTTTAAATTAATGTGTGTAATCTTGTGCTTCTGATACCA 969
```

Mon Nov 18 08:42:02 2002

us-09-820-003a-1.rst

Page 13

Db	361	AAACCTTGTGTGTGCTGGTTTATATAAATATAGTGTATATCCGTCGTTCCGTGACACAA	420
QY	970	GACGTCTTCCCGTGGTGGTTAGATAAATATATTTTGGTGGTGA--TGTTATATTTGGCATGTT	1027
Db	421	GACGTCTTCCCGTGGTGGTTAGATAAATATTTTGGTGGTGA--TGTTATATTTGGCATGTT	480
QY	1028	TAGATGTACAGTTTACGTTCTCTGTGAAGTAAAGAGTACACGATTTTGTATACAAACAGACAA	1087
Db	481	TGAGATGACAGTTTATGCTCTGTGAAGTAAAGAGTACACGATTTGATATCAACAGACAA	540
QY	1088	GAGATGTCTGTCACTTTTGGTACATGATTAAGTTAGTGAAGATGTTATATGTAAAGATCGATT	1147
Db	541	GAGATGTCTGTCACTTTTGGTACATGATTAAGTTAGTGAAGATGTTATATGTAAAGATCGATT	600
QY	1148	TGCTATGTTCTTCTGTGAGATTTTAAATGAGAA--AGATTACATATCGATTAAATAGTGT	1206
Db	601	TGCTATGTTCTTCTGTGAGATTTTAAATGAGAAAGATTAACATATCGATTAAATAGTGT	660
QY	1207	CTTTCATATCTGCAATATTAATTTGGCTGAGATAATATGTAATTTTGGACACACTATGTA	1266
Db	661	CTTTCATATCTGCAATATTAATTTGGCTGAGATAATATGTAATTTTGGACACACTATGTA	720
QY	1267	ACCAAAACACATGAG 1281	
Db	721	ACCAAAACACATGAG 735	

Search completed: November 17, 2002, 13:01:48
Job time : 1946 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 11:38:17 ; Search time 66 Seconds
(Without alignments) 349.278 Million cell updates/sec

Title: US-09-820-003a-2

Perfect score: 886
Sequence: 1 MSSMPKDYFLFKLLIGDS.....EKSNNKIOSTPVKQSGGCC 173

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
1: A_Geneseq.101002.*
2: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SID22/gcgcdata/geneseq/geneseqp-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	860	97.1	205	21	AA34843
2	860	97.1	205	21	AA34844
3	860	97.1	205	21	AA34845
4	768.5	86.7	221	22	AA013525
5	768.5	86.7	221	22	AA013525
6	768.5	86.7	221	22	AA013525
7	768.5	86.7	221	22	AA013525
8	768.5	86.7	221	22	AA013525
9	768.5	86.7	221	22	AA013525
10	768.5	86.7	221	22	AA013525

11	619.5	69.9	258	21	AA30498	Arabidopsis thalia
12	618.5	69.8	203	21	AA30499	Arabidopsis thalia
13	612	69.1	218	21	AA30710	Arabidopsis thalia
14	612	69.1	221	21	AA30709	Arabidopsis thalia
15	612	69.1	254	21	AA30533	Arabidopsis thalia
16	580	65.5	197	21	AA34820	Arabidopsis thalia
17	580	65.5	218	21	AA34820	Arabidopsis thalia
18	442	49.9	216	21	AA30806	Arabidopsis thalia
19	442	49.9	216	21	AA30806	Arabidopsis thalia
20	429.5	48.5	224	21	AA30806	Arabidopsis thalia
21	429.5	48.5	224	21	AA30806	Arabidopsis thalia
22	428.5	48.4	216	21	AA30806	Arabidopsis thalia
23	428.5	48.4	216	21	AA30806	Arabidopsis thalia
24	428.5	48.4	253	21	AA30806	Arabidopsis thalia
25	428.5	48.4	254	21	AA30806	Arabidopsis thalia
26	419	47.3	213	22	AB11916	Human Rab10 protein
27	417	47.1	221	23	AB11916	Human Rab10 protein
28	414	46.7	207	22	AB11916	Human Rab10 protein
29	414	46.7	207	22	AB11916	Human Rab10 protein
30	409.5	46.2	335	22	AB11916	Human Rab10 protein
31	408	46.0	207	22	AB11916	Human Rab10 protein
32	405	45.7	163	21	AA31848	Human Rab10 protein
33	403.5	45.5	200	21	AA31848	Human Rab10 protein
34	403.5	45.5	200	21	AA31848	Human Rab10 protein
35	403.5	45.5	200	21	AA31848	Human Rab10 protein
36	403.5	45.5	200	21	AA31848	Human Rab10 protein
37	403	45.1	141	21	AA31848	Human Rab10 protein
38	399.5	45.1	218	22	AB10726	Human Rab10 protein
39	393	44.4	201	21	AA30982	Human Rab10 protein
40	393	44.4	204	22	AB10760	Human Rab10 protein
41	390.5	44.1	209	22	AB10760	Human Rab10 protein
42	390.5	44.1	246	21	AA35816	Human Rab10 protein
43	387	43.7	199	21	AA30980	Human Rab10 protein
44	387	43.7	199	21	AA30980	Human Rab10 protein
45	387	43.7	199	21	AA30980	Human Rab10 protein

ALIGNMENTS

RESULT 1	AA34843	standard; Protein; 205 AA.
ID	AA34843	
AC	AA34843	
XX	26-JAN-2001	(first entry)
XX	Gene 44 human secreted protein homologous amino acid sequence #131.	
XX	Human; secreted protein; diagnosis: cytostatic; immunosuppressive;	
XX	nocturnal; neuroprotective; antiviral; antiallergic; hepatotropic;	
XX	antidiabetic; antiinflammatory; antitumor; antiparasitic; anticonvulsant;	
XX	antibacterial; antifungal; antiparasitic; antiparasitic; antiparasitic;	
XX	cancer; immune disorder; cardiovascular disorder; wound healing;	
XX	neurological disease; infectious disease; chromosome identification.	
XX	Homo sapiens.	
XX	W0200058356-A1.	
XX	05-OCT-2000.	
XX	22-MAR-2000; 2000WO-US07353.	
XX	26-MAR-1999; 99US-0126511.	
XX	17-DEC-1999; 99US-0172413.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
XX	Rosen CA, Ruben SM, Komatsoulis G;	
XX	WPI: 2000-594639/56.	

XX Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
PS Disclosure; Page 415-416; 425pp; English.
XX
CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
CC AAB34852 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neurotropic; hepatotropic;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.
XX
SQ Sequence 205 AA;
Query Match 97.1%; Score 860; DB 21; Length 205;
Best Local Similarity 84.4%; Pred. No. 7.6e-79;
Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
QY 1 MSSNMPEYDYLFLKLLIGDSGVKSKCLLRPADYTESYSTIGVDFKRTIELDGKTI 60
DB 1 MSSNMPEYDYLFLKLLIGDSGVKSKCLLRPADYTESYSTIGVDFKRTIELDGKTI 60
QY 61 KLOI-----ESFNVMKQLOEIDRYASENVKL 88
DB 61 KLOIWDTAGGERFRTITSSYRGAGHIIYVDYDQESFNVMKQLOEIDRYASENVKL 120
QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKRMGPGA 148
DB 121 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKRMGPGA 180
QY 149 TAGGAESNVKRIQSTPVKQSGGCC 173
DB 181 TAGGAESNVKRIQSTPVKQSGGCC 205
RESULT 2
AAB34844
ID AAB34844 standard; Protein: 205 AA.
XX
AC AAB34844;
XX
DT 26-JAN-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 44 SEQ ID NO:132.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KW cancer; immune disorder; cardiovascular disorder; wound healing;
KW neurological disease; infectious disease; chromosome identification.
XX
OS Homo sapiens.
XX

PN WO200058356-A1.
XX
XX 05-OCT-2000.
XX
XX 22-MAR-2000; 2000WO-US07535.
XX
XX 26-MAR-1999; 99US-0126511.
XX 17-DEC-1999; 99US-0172413.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-594639/56.
DR
XX
PT Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
PS Disclosure; Page 416-417; 425pp; English.
XX
CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
CC AAB34852 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders,
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.
XX
SQ Sequence 205 AA;
Query Match 97.1%; Score 860; DB 21; Length 205;
Best Local Similarity 84.4%; Pred. No. 7.6e-79;
Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;
QY 1 MSSNMPEYDYLFLKLLIGDSGVKSKCLLRPADYTESYSTIGVDFKRTIELDGKTI 60
DB 1 MSSNMPEYDYLFLKLLIGDSGVKSKCLLRPADYTESYSTIGVDFKRTIELDGKTI 60
QY 61 KLOI-----ESFNVMKQLOEIDRYASENVKL 88
DB 61 KLOIWDTAGGERFRTITSSYRGAGHIIYVDYDQESFNVMKQLOEIDRYASENVKL 120
QY 89 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKRMGPGA 148
DB 121 LVGNKCDLTTKKVVDYTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKRMGPGA 180
QY 149 TAGGAESNVKRIQSTPVKQSGGCC 173
DB 181 TAGGAESNVKRIQSTPVKQSGGCC 205
RESULT 3
AAB34816
ID AAB34816 standard; Protein: 198 AA.
XX
AC AAB34816;

XX 26-JAN-2001 (first entry)
DT Human secreted protein sequence encoded by gene 44 SEQ ID NO:104.
DE
XX
XX Human: secreted protein; diagnosis; cytostatic; immunosuppressive;
KW neotropic; neuroprotective; antiviral; antidiabetic; hepatotropic;
KW antidiabetic; antineoplastic; antitumor; vulnere; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KW cancer; immune disorder; cardiovascular disorder; wound healing;
KW neurological disease; infectious disease; chromosome identification.
XX
OS Homo sapiens.
XX
PN WO200058356-A1.
XX
PD 05-OCT-2000.
XX
PF 22-MAR-2000; 2000MO-US07535.
XX
PR 26-MAR-1999; 99US-0126511.
PR 17-DEC-1999; 99US-0172413.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI: 2000-594639/56.
DR N-PSDB: AAC60009.
XX
PT Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
PS Claim 1; Page 393-394; 425pp; English.
XX
XX The polynucleotide sequences given in AAC59966 to AAC60015 encode the
CC human secreted proteins given in AAB34773 to AAB34823 to
CC AAB34852 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral;
CC vulnerary; anticonvulsant; antidiabetic; antineoplastic; antitumor;
CC cardiant; polynucleotide; antibacterial; antifungal; antiparasitic; and
CC preventing, treating or ameliorating a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. The polynucleotides are
CC useful for chromosome identification. They are also useful as probes for
CC diagnosing a disorder related to the female reproductive system,
CC particularly breast and/or ovary cancer. They are also useful in the gene
CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
CC agonists and antagonists from the present invention are useful in the
CC diagnosis, treatment and prevention of cancer, immune disorders, and
CC cardiovascular disorders, wound healing, neurological diseases and
CC infectious disease. AAC59966 to AAC59965 and AAB34772 represents sequence
CC used in the exemplification of the present invention.
XX
SQ Sequence 198 AA:
Query Match 92.2%; Score 817; DB 21; Length 198;
Best Local Similarity 83.8%; Pred. No. 1.6e-74;
Matches 165; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

OY 97 TTKKVDYTTAKEFADSLGIPFLETSKAKNATNVEOSFMTMAEIKRMPGATAGAKRS 156
DB 121 TTKKVDYTTAKEFADSLGIPFLETSKAKNATNVEOSFMTMAEIKRMPGATAGAKRS 180
OY 157 NVKIOSPPVKOSGGGCC 173
DB 181 NVKIOSPPVKOSGGGCC 197
RESULT 4
ID AA013525
XX AA013525 standard; Protein; 221 AA.
XX
AC AA013525;
XX
PD 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 27417.
XX
XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001MO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSPO INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-514838/56.
DR N-PSDB: AA193456.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 27417; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 221 AA:
Query Match 86.7%; Score 768.5; DB 22; Length 221;
Best Local Similarity 76.5%; Pred. No. 1.5e-69;
Matches 156; Conservative 8; Mismatches 7; Indels 33; Gaps 2;

OY 2 SSMPPEYDULFKLLIGDSGVGKSCLLAFADDTYTESYSTIGVDFKRIETEDGKTIK 61
DB 19 AANPEYDULFKLLIGDSGVGKSCLLAFADDTYTESYSTIGVDFKRIETEDGKTIK 78

QY 62 LQI-----ESFNNVKQWLQIEDRYASENVKLL 89
 |||
 Db 79 LQIWDTAGQERFRITSSYRGAGHIIVVDYDTQESYANVKQWLQIEDRYASENVKLL 138
 QY 90 VGNKCDLTTKKVVDYTTAKEFADSLGIPLETSAKNATNVQSFMTMAEIKRMRGCGAT 149
 |||
 Db 139 VGNKSDLTTRKKVVDNTTAKFADSLGIPLETSAKNATNVQSFMTMAEIKRMRGCGAA 198
 QY 150 AGAEEKSNVKRQSTPVPKQSGGCC 173
 |||
 Db 199 SGG-ERPLKIDSTPVPKPAAGGCC 221

RESULT 5
 AAB58758
 ID AAB58758 standard; Protein: 224 AA.
 XX
 AC AAB58758;
 XX
 D7 27-MAR-2001 (first entry)
 XX
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 466.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antitumor; vulnereary; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.
 KW
 XX Homo sapiens.
 OS
 PN WO200055173-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05881.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM;
 XX
 DR WPI: 2000-611515/58.
 DR N-PADB; AAF21661.
 XX
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -
 XX
 PS Claim 11; Page 895-896; 1299pp; English.
 XX
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antitumor; vulnereary; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid

CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 SQ Sequence 224 AA;
 Query Match 86.7%; Score 768.5; DB 21; Length 224;
 Best Local Similarity 76.5%; Pred. No. 1.5e-69;
 Matches 156; Conservative 8; Mismatches 7; Indels 33; Gaps 2;

QY 2 SSANPEYDYLKLLLLIDSGVGSCLLRFADPTTYSYSTIVDPKIRIIEIDGKTK 61
 |||
 Db 22 AANPEYDYLKLLLLIDSGVGSCLLRFADPTTYSYSTIVDPKIRIIEIDGKTK 81
 QY 62 LQI-----ESFNNVKQWLQIEDRYASENVKLL 89
 |||
 Db 82 LQIWDTAGQERFRITSSYRGAGHIIVVDYDTQESYANVKQWLQIEDRYASENVKLL 141
 QY 90 VGNKCDLTTKKVVDYTTAKEFADSLGIPLETSAKNATNVQSFMTMAEIKRMRGCGAT 149
 |||
 Db 142 VGNKSDLTTRKKVVDNTTAKFADSLGIPLETSAKNATNVQSFMTMAEIKRMRGCGAA 201
 QY 150 AGAEEKSNVKRQSTPVPKQSGGCC 173
 |||
 Db 202 SGG-ERPLKIDSTPVPKPAAGGCC 224

RESULT 6
 AAY00919
 ID AAY00919 standard; Protein: 201 AA.
 XX
 AC AAY00919;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE Human Rab protein, RABP-2, protein sequence.
 XX
 KW Rab protein; RABP-1; RABP-2; RABP-3; human; vesicle trafficking; cancer;
 KW cell differentiation; apoptosis; immunodeficiency; cell proliferation;
 KW neurodegenerative disease; myelodysplastic syndrome; wasting disease;
 KW toxin-induced disease; infection; genetic defect; diagnosis; therapy.
 XX
 OS Homo sapiens.
 PN WO9909182-A2.
 XX
 PD 25-FEB-1999.
 XX
 PF 17-AUG-1998; 98WO-US16983.
 XX
 PR 21-AUG-1997; 97US-0916901.
 XX
 PA (INCY-) INCYTE PHARM INC.
 PI Corley NC, Hillman JL, Lal P, Shah P;
 XX
 DR WPI: 1999-181042/15.
 DR N-PADB; AAX27232.
 XX
 PT New purified human Rab proteins - used to develop products for
 PT treating e.g. AIDS, immunodeficiencies, neurodegenerative diseases,
 PT cancer, inflammation or autoimmune diseases.
 XX
 PS Claim 23; Fig 2; 94pp; English.
 XX
 CC This sequence is a human Rab protein of the invention, designated
 CC RABP-2. The Rab proteins, RABP-1, RABP-2 and RABP-3 are involved in
 CC vesicle trafficking, cell function, and cell differentiation. The RABP
 CC polypeptides, DNAs and agonists can be used to prevent or treat a
 CC disorder associated with an increase in apoptosis, e.g. infectious or
 CC genetic immunodeficiencies, neurodegenerative diseases such as
 CC Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis,
 CC retinitis pigmentosa, and cerebellar degeneration, myelodysplastic

CC syndromes such as aplastic anaemia, ischaemic injuries such as myocardial
 CC infarction, stroke, and reperfusion injury, toxin-induced diseases such
 CC as alcohol-induced liver damage, cirrhosis, and lathyrism, wasting
 CC diseases such as cachexia, viral infections, and osteoporosis. They can
 CC also be used to stimulate cell proliferation for use in transplantation
 CC or to produce cells to fight an infection or a cancer or to correct a
 CC genetic defect in a disease such as sickle cell beta thalassemia, cystic
 CC fibrosis or Huntington's chorea. Antagonists can be used to prevent or
 CC treat a disorder associated with cell proliferation e.g. cancers or
 CC inflammation, e.g. Addison's disease, adult respiratory distress
 CC syndrome, allergies, asthma, atherosclerosis, bronchitis, cholecystitis,
 CC Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis,
 CC gout, Graves' disease, hyperosinophilia, irritable bowel syndrome, lupus
 CC erythematosus, multiple sclerosis, myasthenia gravis, inflammation,
 CC osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid
 CC arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis,
 CC complications of cancer, haemodialysis, extracorporeal circulation,
 CC infections and trauma.

XX Sequence 201 AA:
 SQ

Query Match 86.5%; Score 766.5; DB 20; Length 201;
 Best local similarity 77.2%; Pred. No. 2,1e-69;
 Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;

QY 4 MNPEVDYIFKLLIGDSGVKSCLLRFADDTYESTYSTGVPRKRTEDGKTIKQ 63
 DB 1 MNPEVDYIFKLLIGDSGVKSCLLRFADDTYESTYSTGVPRKRTEDGKTIKQ 60
 QY 64 I-----ESPNNVKOMLOEDIRYASENVKLLVG 91
 DB 61 IMDTAGEERFTITSSYRGAGIIYVDYTDQESYANKOMLOEDIRYASENVKLLVG 120
 QY 92 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNAITNVEOSFTMAAEIKRRMGPGATAG 151
 DB 121 NKSDLTTKKVVDYTTAKFADSLGIPLETSAKNAITNVEOSFTMAAEIKRRMGPGASG 180
 QY 152 GAERSNVKTOSTPVKOSGGCC 173
 DB 181 G-ERPNEKIDSTVPKAGGCC 201

RESULT 7
 AAU28024
 ID AAU28024 standard; Protein: 201 AA.
 XX
 AC AAU28024:
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 193.
 XX
 KM Human: secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KM ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KM transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KM amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KM ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KM gut protection; lung; liver fibrosis; immune deficiency; infection;
 KM severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KM multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KM fertility; analgesic; pain; antigen.
 XX
 OS Homo sapiens.
 XX
 PN MO200166689-A2.
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001MO-US04942.
 XX
 RA 07-MAR-2001; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.

PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX
 PA (HVS-) HVS00 INC.
 PI Zhang Y, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Tang QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 DR WPI: 2001-589934/66.
 DR N-PSDB: AAS44924.
 PS
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX
 XX Example 3; SEQ ID No 193; 107pp; English.

CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC treating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, vitamins, minerals, provides
 CC dietary fat, lipid, protein, carbohydrate, storage or elimination of
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.
 XX
 XX Sequence 201 AA:
 SQ

Query Match 86.5%; Score 766.5; DB 22; Length 201;
 Best local similarity 77.2%; Pred. No. 2,1e-69;
 Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;

QY 4 MNPEVDYIFKLLIGDSGVKSCLLRFADDTYESTYSTGVPRKRTEDGKTIKQ 63
 DB 1 MNPEVDYIFKLLIGDSGVKSCLLRFADDTYESTYSTGVPRKRTEDGKTIKQ 60
 QY 64 I-----ESPNNVKOMLOEDIRYASENVKLLVG 91
 DB 61 IMDTAGEERFTITSSYRGAGIIYVDYTDQESYANKOMLOEDIRYASENVKLLVG 120
 QY 92 NKCDLTTKKVVDYTTAKFADSLGIPLETSAKNAITNVEOSFTMAAEIKRRMGPGATAG 151
 DB 121 NKSDLTTKKVVDYTTAKFADSLGIPLETSAKNAITNVEOSFTMAAEIKRRMGPGASG 180
 QY 152 GAERSNVKTOSTPVKOSGGCC 173
 DB 181 G-ERPNEKIDSTVPKAGGCC 201

Db 181 G-ERPMLKIDSTPVKPGAGGCC 201

RESULT 8

ABB59808
ID ABB59808 standard; Protein: 205 AA.

XX ABB59808;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 6216.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001MO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li FWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL03911.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 6216; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 205 AA;

Query Match 77.1%; Score 683; DB 22; Length 205;

Best Local Similarity 69.4%; Pred. No. 6e-61;

Matches 143; Conservative 12; Mismatches 17; Indels 34; Gaps 3;

QY 1 MSSNNPEVDYLFKLLIGDSGVKSCLLRRADTYTTSYSTIGVDKRTIELDGTI 60

Db 1 MSSVNPEDYLFKLLIGDSGVKSCLLRRADTYTTSYSTIGVDKRTIELDGTI 60

QY 61 KLQI-----ESFNWVKWLOEIDRYASENVKL 88

Db 61 KLQIWDITAGQERFRITTSYRGAGIIVVDCIDQESFNWVKWLEIETRYACENVKL 120

QY 89 LVGNKCDLITRKVVYDTTAKFADSLGIPLETSAKNATNVEQSFMTAAEIKRMGPGA 148

Db 121 LVGNKSDLTITRKVVYDHTTAAEYLAQIGIPLETSAKSNATNVEQAFMTAAEIKRNVGPS 180

QY 149 TAGAEEKSNVAI-OSTPVKQSGGCC 173

Db 181 SATD-NASKYKIDQGRPVENKTRSGCC 205

RESULT 9

AAG07763
ID AAG07763 standard; Protein: 202 AA.

XX AAG07763;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 5040.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126254.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148568.
PR 16-AUG-1999; 99US-0148568.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158322.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 71.9%; Score 637; DB 21; Length 202;
Best Local Similarity 65.3%; Pred No 2 6e-56;
Matches 122; Conservatve 11; Mismatches 25; Indels 34; Gaps 3;

QY 4 MNEPYIKLLIGDSGCKSLLRPRADDTYESISTIGVDKIKITLLOCKIRIQ 63
DB 1 MNEPYIKLLIGDSGCKSLLRPRADDTYESISTIGVDKIKITLLOCKIRIQ 60
QY 64 I-----ESPNNKQWLOEIDRYASEVNNKLLVG 91
DB 61 IMDPAQERFTIISYRKAGHGIYTVDTDESFPNNKQWLEIDRYASEVNNKLLVG 120
QY 92 NCDLTKKVVDTTAAKEPADSIGIPLETISAKNATNVEQSFMTAAEIKRRMGPGATAG 151
|||||:|||| ||| ||| |||||||||||||||:|||| |||


```
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149476.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158038.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158309.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

```
Query Match          71.4% Score 633; DB 21; Length 202;
Best Local Similarity 65.8%; Pred No. 6; 7e-56;
Matches 133; Conservative 10; Mismatches 25; Indels 34; Gaps 3;
```

QY 4 MNPEYDLFFALLIDGSGVSCLLRFADDTYTSYSTIGDVFIRIIEIDGKTKIKQ 63
DB 1 MNPEYDLFFALLIDGSGVSCLLRFADDTYTSYSTIGDVFIRIIEIDGKTKIKQ 60
QY 64 I-----ESFNNVKOMIQEIDRYASENNKLLVG 91

```
DB 61 IWDPAQGEFRTITSSYYRGAGSIITYVDYDLESEFNNVKOMIQEIDRYASENNKLLVG 120
QY 92 NKCDLTKRKVDYTTAAKEFADSLGIFPLETSAKNATNVESQFMTAAEIKRKPGPATAG 151
DB 121 NKNDLTSOKVSTEFARAFADDELGIPLFETSANKNATNVEBAMATTAIKTRNA-SOPAG 179
QY 152 GAESKNVKIQTSPYKQSGGGCC 173
DB 180 GAKPPTVQIRGQPVNOC-SGCC 200

RESULT 11
AAG30498
ID AAG30498 standard; Protein; 258 AA.
AC AAG30498;
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 36472.
XX DE Protein identification: signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX Arabidopsis thaliana.
XX EPI033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF XX
XX 25-FEB-1999; 99US-0121825.
PR XX 05-MAR-1999; 99US-0123180.
PR XX 09-MAR-1999; 99US-0123548.
PR XX 23-MAR-1999; 99US-0125788.
PR XX 25-MAR-1999; 99US-0126264.
PR XX 29-MAR-1999; 99US-0126785.
PR XX 01-APR-1999; 99US-0127462.
PR XX 06-APR-1999; 99US-0128234.
PR XX 08-APR-1999; 99US-0128714.
PR XX 16-APR-1999; 99US-0129845.
PR XX 19-APR-1999; 99US-0130077.
PR XX 21-APR-1999; 99US-0130449.
PR XX 23-APR-1999; 99US-0130510.
PR XX 23-APR-1999; 99US-0130891.
PR XX 28-APR-1999; 99US-0131449.
PR XX 30-APR-1999; 99US-0132048.
PR XX 04-MAY-1999; 99US-0132407.
PR XX 05-MAY-1999; 99US-0132484.
PR XX 06-MAY-1999; 99US-0132485.
PR XX 06-MAY-1999; 99US-0132486.
PR XX 07-MAY-1999; 99US-0132863.
PR XX 11-MAY-1999; 99US-0134256.
PR XX 14-MAY-1999; 99US-0134218.
PR XX 14-MAY-1999; 99US-0134421.
PR XX 14-MAY-1999; 99US-0134470.
PR XX 18-MAY-1999; 99US-0134768.
PR XX 19-MAY-1999; 99US-0134841.
PR XX 20-MAY-1999; 99US-0135124.
PR XX 21-MAY-1999; 99US-0135353.
PR XX 24-MAY-1999; 99US-0135629.
PR XX 25-MAY-1999; 99US-0136021.
PR XX 27-MAY-1999; 99US-0136982.
PR XX 28-MAY-1999; 99US-0136782.
PR XX 01-JUN-1999; 99US-0137722.
PR XX 03-JUN-1999; 99US-0137528.
PR XX 04-JUN-1999; 99US-0137502.
PR XX 07-JUN-1999; 99US-0137724.
```

PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140685.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142350.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0143542.
 PR 13-JUL-1999; 99US-0143624.
 PR 14-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 19-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 69.94; Score 619.5; DB 21; Length 258;
 Best Local Similarity 61.64; Pred. No. 2.1e-54;
 Matches 125; Conservative 15; Mismatches 30; Indels 33; Gaps 2;

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149475.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151060.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160960.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.8%; Score 618.5; DB 21; Length 203;
Best Local Similarity 61.9%; Pred. No. 2e-54;
Matches 125; Conservative 14; Mismatches 30; Indels 33; Gaps 2;

QY 4 MNPEYDYLFKLLIGDSVGSCKLLRPADTYTESYSTGVDPKRTIILDGKTKLQ 63
Db 1 MNPEYDYLFKLLIGDSVGSCKLLRPADSYVESYSTIGVDFKRTVEQDKTKLQ 60
QY 64 I-----ESFNNVKOMLOEDRASEYNNKLLYG 91
Db 61 IWDTAGOERFRTTSSYRGAHGIIIVDYDESFNNVKOMLSEIDRYASDNKLLYG 120
QY 92 NKCDLTKRKVYDTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIRKRMGPATAG 151
Db 121 NKSDLTENRAIPYETAKAFADIEIGIFWETSAKDATTNEQAFMMSASIKRMSAPAGN 180
QY 152 GAESKNVKIQTSTPVKQSGGCC 173
Db 181 NARPTVOIRGQPYAOK-NGCC 201

RESULT 13
AAG30710
ID AAG30710 standard; Protein; 218 AA.
XX
AC AAG30710;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36764.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128274.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

Mon Nov 18 08:42:02 2002

us-09-820-003a-2.rag

Page 13

PR	20-MAY-1999;	9905-0135124.
PR	21-MAY-1999;	9905-0135333.
PR	24-MAY-1999;	9905-0135629.
PR	25-MAY-1999;	9905-0136021.
PR	27-MAY-1999;	9905-0136392.
PR	28-MAY-1999;	9905-0136782.
PR	01-JUN-1999;	9905-0137222.
PR	03-JUN-1999;	9905-0137528.
PR	04-JUN-1999;	9905-0137502.
PR	07-JUN-1999;	9905-0137724.
PR	08-JUN-1999;	9905-0138094.
PR	10-JUN-1999;	9905-0138540.
PR	10-JUN-1999;	9905-0138847.
PR	14-JUN-1999;	9905-0139119.
PR	16-JUN-1999;	9905-0139452.
PR	17-JUN-1999;	9905-0139453.
PR	18-JUN-1999;	9905-0139454.
PR	18-JUN-1999;	9905-0139455.
PR	18-JUN-1999;	9905-0139456.
PR	18-JUN-1999;	9905-0139457.
PR	18-JUN-1999;	9905-0139458.
PR	18-JUN-1999;	9905-0139459.
PR	18-JUN-1999;	9905-0139460.
PR	18-JUN-1999;	9905-0139461.
PR	18-JUN-1999;	9905-0139462.
PR	18-JUN-1999;	9905-0139463.
PR	18-JUN-1999;	9905-0139750.
PR	18-JUN-1999;	9905-0139763.
PR	21-JUN-1999;	9905-0139817.
PR	22-JUN-1999;	9905-0139899.
PR	23-JUN-1999;	9905-0140353.
PR	23-JUN-1999;	9905-0140354.
PR	24-JUN-1999;	9905-0140695.
PR	28-JUN-1999;	9905-0140823.
PR	29-JUN-1999;	9905-0140991.
PR	30-JUN-1999;	9905-0141287.
PR	01-JUL-1999;	9905-0141842.
PR	01-JUL-1999;	9905-0142154.
PR	02-JUL-1999;	9905-0142055.
PR	06-JUL-1999;	9905-0142390.
PR	08-JUL-1999;	9905-0142803.
PR	09-JUL-1999;	9905-0142820.
PR	12-JUL-1999;	9905-0142870.
PR	13-JUL-1999;	9905-0143342.
PR	14-JUL-1999;	9905-01433624.
PR	15-JUL-1999;	9905-0144085.
PR	16-JUL-1999;	9905-0144085.
PR	19-JUL-1999;	9905-0144326.
PR	19-JUL-1999;	9905-0144332.
PR	19-JUL-1999;	9905-0144337.
PR	19-JUL-1999;	9905-0144333.
PR	19-JUL-1999;	9905-0144334.
PR	19-JUL-1999;	9905-0144335.
PR	20-JUL-1999;	9905-0144352.
PR	20-JUL-1999;	9905-0144632.
PR	20-JUL-1999;	9905-0144684.
PR	21-JUL-1999;	9905-0144684.
PR	21-JUL-1999;	9905-0145086.
PR	22-JUL-1999;	9905-0145088.
PR	22-JUL-1999;	9905-0145085.
PR	22-JUL-1999;	9905-0145087.
PR	22-JUL-1999;	9905-0145089.
PR	23-JUL-1999;	9905-0145192.
PR	23-JUL-1999;	9905-0145145.
PR	23-JUL-1999;	9905-0145224.
PR	26-JUL-1999;	9905-0145276.
PR	27-JUL-1999;	9905-0145913.
PR	27-JUL-1999;	9905-0145918.
PR	28-JUL-1999;	9905-0145951.
PR	02-AUG-1999;	9905-0146388.
PR	02-AUG-1999;	9905-0146389.
PR	03-AUG-1999;	9905-0147038.
PR	04-AUG-1999;	9905-0147204.
PR	04-AUG-1999;	9905-0147302.
PR	05-AUG-1999;	9905-0147192.
PR	05-AUG-1999;	9905-0147260.
PR	06-AUG-1999;	9905-0147303.
PR	06-AUG-1999;	9905-0147416.
PR	09-AUG-1999;	9905-0147935.
PR	10-AUG-1999;	9905-0148171.
PR	11-AUG-1999;	9905-0148319.
PR	13-AUG-1999;	9905-0148319.
PR	13-AUG-1999;	9905-0148565.
PR	16-AUG-1999;	9905-0149368.
PR	17-AUG-1999;	9905-0149175.
PR	18-AUG-1999;	9905-0149246.
PR	20-AUG-1999;	9905-0149722.
PR	20-AUG-1999;	9905-0149929.
PR	23-AUG-1999;	9905-0149930.
PR	23-AUG-1999;	9905-0149930.
PR	25-AUG-1999;	9905-0150566.
PR	26-AUG-1999;	9905-0150884.
PR	27-AUG-1999;	9905-0151065.
PR	27-AUG-1999;	9905-0151065.
PR	27-AUG-1999;	9905-0151080.
PR	30-AUG-1999;	9905-0151303.
PR	31-AUG-1999;	9905-0151348.
PR	01-SEP-1999;	9905-0151930.
PR	07-SEP-1999;	9905-0152363.
PR	10-SEP-1999;	9905-01530

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.18; Score 612; DB 21; Length 218;
Best Local Similarity 65.08; Pred. No. 9.8e-54;
Matches 128; Conservative 11; Mismatches 24; Indels 34; Gaps 3;

OY 9 DYLFKLLIGSGVKSCLLRFADDDTSEYISITGVDFKRTIELDKTITKIQI---- 64
D 22 DYLFKLLIGSGVKSCLLRFADDDTSEYISITGVDFKRTIELDKTITKIQI 81
OY 65 -----ESFNWVKQLOEIDRYASENVNKLIVGKCDL 96
D 82 GQERFRITTSYRGAGHIIYTDVTDLESFNWVKQLOEIDRYASENVNKLIVGKCDL 141
OY 97 TTKRVVDYTTAKERPADSIGIPLETSAKATNVEQSEFTMAAEIKRKGPGATAGAEKS 156
D 142 TSQKRVSTETAKAFADDELGIPELETSAKATNVEEAFVAMTAATKTRMA-SQPAGSKRP 200
OY 157 NKKIOSFPVRKSGGCC 173
D 201 TVQIRGPFVNOQ-SGCC 216

RESULT 14

AAG30709
ID AAG30709 standard; Protein: 221 AA.

AC AAG30709;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 36763.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148564.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 69.1%; Score 612; DB 21; Length 221;
Best Local Similarity 65.0%; Pred. No. 1e-53;
Matches 18; Conservative 11; Mismatches 24; Indels 34; Gaps 3;

QY 9 DYLFKLLIGDSGVKSCLLRFADPTYESYSTIGVDKFKIRITLDTGKTIKLT---- 64
DB 25 DYLFKLLIGDSGVKSCLLRFADDSYLDYSITIGVDKFKIRTEQDKTIKLTWDTA 84

QY 65 -----ESFNMYKQLEIDIRVASEVYKLLVGNKKDL 96
DB 85 GDRFRFTTSSYYTGAGAGITVYDVTDLSEFNMYKQLEIDIRVASEVYKLLVGNKKDL 144

QY 97 TTKKVVDTYTAKEFADSLGIPELETSANNAATNVEQSFMTAAEIKRMGPGATAGAEKS 156
DB 145 TSQKVVSTETAKAFADSLGIPELETSANNAATNVEAFMTAATKTRMA-SQPAAGSKPP 203

QY 157 NVRIGSTPYKQSGGCC 173
DB 204 TVQIRQPFVNOQ-SGCC 219

RESULT 15
AAG45323 standard; Protein; 254 AA.
ID AAG45323;
AC AAG45323;
XX
XX
DT 18-OCT-2000 (first entry)
XX
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56887.
XX
XX
KM Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139432.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144351.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 13:04:12 : Search time 20 Seconds
(without alignments)
130.277 Million cell updates/sec

Title: US-09-820-003a-2
Perfect score: 886
Sequence: 1 MSNNPEYDYLFKLLIGDS.....EKSNVKIOSTPYKOSGGGCC 173

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5
Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCU_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	886	100.0	173	US-09-820-003a-2	Sequence 2, Appl1
2	866	97.1	222	US-09-820-003a-4	Sequence 4, Appl1
3	766.5	86.3	201	US-09-567-736-3	Sequence 3, Appl1
4	766.5	85.8	201	US-09-567-736-3	Sequence 8, Appl1
5	414	46.7	207	US-09-794-257-8	Sequence 5, Appl1
6	403.5	44.1	218	US-09-822-860-5	Sequence 5, Appl1
7	390.5	43.5	218	US-09-925-300-1571	Sequence 1571, App
8	371	41.9	162	US-09-834-765-766	Sequence 534, App
9	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
10	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
11	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
12	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
13	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
14	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
15	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
16	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
17	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
18	362.5	40.9	198	US-09-925-302-534	Sequence 16, App
19	292.5	33.0	191	US-09-794-257-14	Sequence 14, Appl1

20	292.5	33.0	191	US-10-051-986-3	Sequence 3, Appl1
21	292.5	33.0	312	US-09-925-302-783	Sequence 783, App
22	292	33.0	59	US-09-864-761-35038	Sequence 35038, A
23	267	30.1	216	US-09-945-173-10	Sequence 10, Appl1
24	258	29.1	239	US-09-925-301-1077	Sequence 1077, App
25	255	28.1	217	US-09-988-974-3	Sequence 5, Appl1
26	254	28.7	213	US-09-794-257-3	Sequence 8, Appl1
27	254	28.7	213	US-09-988-974-8	Sequence 2, Appl1
28	237.5	26.8	208	US-10-067-813-16	Sequence 45, Appl1
29	235.5	26.6	832	US-09-834-765-45	Sequence 4, Appl1
30	235	26.5	217	US-09-925-300-1364	Sequence 1364, App
31	219.5	24.8	183	US-09-765-228A-26	Sequence 26, Appl1
32	219.5	24.8	183	US-09-765-228A-28	Sequence 28, Appl1
33	219.5	24.8	189	US-10-104-484-4	Sequence 2, Appl1
34	216.5	24.4	189	US-10-104-484-2	Sequence 16, Appl1
35	214.5	24.2	189	US-10-067-813-16	Sequence 4, Appl1
36	212	23.9	9	US-09-850-373-4	Sequence 5, Appl1
37	212	23.9	624	US-09-834-765-762	Sequence 762, App
38	212	23.8	144	US-09-972-529-7	Sequence 7, Appl1
39	210.5	23.8	201	US-09-822-860-2	Sequence 2, Appl1
40	206	23.3	201	US-09-988-974-9	Sequence 5, Appl1
41	206	23.3	201	US-09-988-974-9	Sequence 9, Appl1
42	206	23.3	209	US-09-864-761-42996	Sequence 42996, A
43	206	23.1	168	US-09-834-765-765	Sequence 765, App
44	205	23.1	208	US-09-925-302-629	Sequence 629, App
45	200	22.6	208	US-09-925-302-629	Sequence 629, App

ALIGNMENTS

RESULT 1
US-09-820-003a-2
Sequence 2, Application US/09820003a
Patent No. US20020142382A1
GENERAL INFORMATION:
APPLICANT: MERCKUOY, Gennady et al.
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
FILE REFERENCE: CLO01196
CURRENT FILING DATE: 2001-03-29
CURRENT APPLICATION NUMBER: US/09/820,003A
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO. 2
LENGTH: 173
TYPE: PRT
ORGANISM: Homo sapien
US-09-820-003a-2

Query Match 100.0% Score 886 DB 10 Length 173:
Best Local Similarity 100.0% Pred. No. 2.5e-80:
Matches 173: Conservative 0 Mismatches 0 Indels 0 Gaps 0:
OY 1 MSNNPEYDYLFKLLIGDSGVKSCLLRFPADDTTSTIGVDFKRTIETDGTI 60
DB 1 MSNNPEYDYLFKLLIGDSGVKSCLLRFPADDTTSTIGVDFKRTIETDGTI 60
OY 61 KLOIESFNNAKQWLOEIDRYASENNKLLVGNKCDLTKKVVDTTAKFADSLGIPLE 120
DB 61 KLOIESFNNAKQWLOEIDRYASENNKLLVGNKCDLTKKVVDTTAKFADSLGIPLE 120
OY 121 TSANATNVEOSEFTMAAEIKRMPGATAGAEKSNVKIOSTPYKOSGGGCC 173
DB 121 TSANATNVEOSEFTMAAEIKRMPGATAGAEKSNVKIOSTPYKOSGGGCC 173
RESULT 2
US-09-820-003a-4
Sequence 4, Application US/09820003a
Patent No. US20020142382A1
GENERAL INFORMATION:

APPLICANT: MERKULOV, Gennady et al.
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: C1001196
CURRENT APPLICATION NUMBER: US/09/820, 003A
CURRENT FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 222
TYPE: PRT
ORGANISM: Homo sapien
US-09-820-003a-4

Query Match 97.1%; Score 860; DB 10; Length 222;
Best Local Similarity 84.4%; Pred. No. 1,2e-77;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MSSMPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYSTIGVDKIRITELDGKTI 60
DB 18 MSSMPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYSTIGVDKIRITELDGKTI 77
QY 61 KLDI-----ESFNNKQWLOEIDRYASENVNKL 88
DB 78 KLDIWDAGERRRTTSSYRGAGHIIYVDYDQESFNNKQWLOEIDRYASENVNKL 137
QY 89 LVGNKCDLTKKVVDTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKKRMGPAG 148
DB 138 LVGNKCDLTKKVVDTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKKRMGPAG 197
QY 149 TAGGAEKSNVKIQSTPVKQSGGCC 173
DB 198 TAGGAEKSNVKIQSTPVKQSGGCC 222

RESULT 3
US-09-967-736-3

Sequence 3, Application US/09967736
Patent No. US20020103340A1

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

LaI, Preeti

Corley, Neil C.

Shah, Purvi

TITLE OF INVENTION: RAB PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/967, 736

FILING DATE: 28-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/154, 602

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0367 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: L1VPT0704

CLONE: 2514506

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-967-736-3

Query Match 86.5%; Score 766.5; DB 10; Length 201;
Best Local Similarity 77.2%; Pred. No. 1.8e-68;

Matches 156; Conservative 6; Mismatches 7; Indels 33; Gaps 2;

QY 4 MNPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYSTIGVDKIRITELDGKTIKQ 63
DB 1 MNPEYDYLFLKLLIGDSGVKSCLLRFADDTYESYSTIGVDKIRITELDGKTIKQ 60
QY 64 I-----ESFNNKQWLOEIDRYASENVNKL 91
DB 61 IWDAGERRRTTSSYRGAGHIIYVDYDQESFNNKQWLOEIDRYASENVNKL 120
QY 92 NKCDLTKKVVDTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKKRMGPAGTAG 151
DB 121 NKSDLTKKVVDTTAKFADSLGIPLETSAKNATNEOSFMTMAEIKKRMGPAGTAG 180
QY 152 GAERKSNVKIQSTPVKQSGGCC 173
DB 181 G-ERPMLKIDSTPVKQSGGCC 201

RESULT 4
US-09-967-736-8

Sequence 8, Application US/09967736
Patent No. US20020103340A1

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

LaI, Preeti

Corley, Neil C.

TITLE OF INVENTION: RAB PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/967, 736

FILING DATE: 28-Sep-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/154, 602

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0367 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 201 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 57006
SEQUENCE DESCRIPTION: SEQ ID NO: 8;
US-09-967-736-8

Query Match 85.8%; Score 760.5; DB 10; Length 201;
Best Local Similarity 76.2%; Pred. No. 6.9e-68;
Matches 154; Conservative 7; Mismatches 8; Indels 33; Gaps 2;

QY 4 NMPEYDLFLKLLIGDSGVGKSCLLRFADDDYTESYSTIGVDFKIRITELDGKTIKQ 63
DB 1 NMPEYDLFLKLLIGDSGVGKSCLLRFADDDYTESYSTIGVDFKIRITELDGKTIKQ 60
QY 64 I-----ESFNNVKOMLOEIDRYASBNVKNLLVG 91
DB 61 IMDRAGGERFRTTAYYRGAMGIMLVYDITNKSFDNIRNIEEHSASDVERMILG 120
QY 92 NKCDLTTKRVVDTTAKFADSLGIPLETSAKNTNVEOSFMTMAEIKRKGAGTAG 151
DB 121 NKSDLTTRKVVDTTAKFADSLGIPLETSAKNTNVEOSFMTMAEIKRKGAGTAG 180
QY 152 GAESKSVKIOSTFVYKSGGCC 173
DB 181 G-ERPNLKDSTFVKSASGGCC 201

RESULT 5
US-09-794-257-8
Sequence 8, Application US/09794257
Patent No. US2002009804A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US2002009804A1el
FILE REFERENCE: 35800/209285
CURRENT APPLICATION NUMBER: US/09/794,257
CURRENT FILING DATE: 2001-02-27
PRIOR APPLICATION NUMBER: 60/185,606
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 8
LENGTH: 207
TYPE: PRT
ORGANISM: homo sapiens
US-09-794-257-8

Query Match 46.7%; Score 414; DB 10; Length 207;
Best Local Similarity 44.2%; Pred. No. 9.7e-34;
Matches 84; Conservative 31; Mismatches 37; Indels 38; Gaps 2;

QY 4 NMPEYDLFLKLLIGDSGVGKSCLLRFADDDYTESYSTIGVDFKIRITELDGKTIKQ 63
DB 1 MAKTYDLFLKLLIGDSGVGKTCLLFRFSDAFNTFTISTIGDFKIRITELDGKTIKQ 60
QY 64 I-----ESFNNVKOMLOEIDRYASBNVKNLLVG 91
DB 61 IMDRAGGERFRTTAYYRGAMGIMLVYDITNKSFDNIRNIEEHSASDVERMILG 120
QY 92 NKCDLTTKRVVDTTAKFADSLGIPLETSAKNTNVEOSFMTMAEIKRKGAGTAG 151
DB 121 NKCDLTTKRVVDTTAKFADSLGIPLETSAKNTNVEOSFMTMAEIKRKGAGTAG 180
QY 146 PGATAGAEK 155
DB 181 NSAGAGPVK 190

RESULT 6
US-09-822-860-5
Sequence 5, Application US/09822860

Patent No. US20020146795A1
GENERAL INFORMATION:
APPLICANT: ZHU, Shaoqing et al.
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
FILE REFERENCE: CLO01214
CURRENT APPLICATION NUMBER: US/09/822,860
CURRENT FILING DATE: 2001-04-02
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5
LENGTH: 190
TYPE: PRT
ORGANISM: Discopyge ommata
US-09-822-860-5

Query Match 46.6%; Score 413; DB 10; Length 190;
Best Local Similarity 42.1%; Pred. No. 1.1e-33;
Matches 60; Conservative 38; Mismatches 40; Indels 32; Gaps 1;

QY 9 DYLFKLLIGDSGVGKSCLLRFADDDYTESYSTIGVDFKIRITELDGKTIKQ---- 64
DB 1 DYLFKLLIGDSGVGKTCLLFRFSDAFNTFTISTIGDFKIRITELDGKTIKQIMDPA 60
QY 65 -----ESFNNVKOMLOEIDRYASBNVKNLLVGKCKDL 96
DB 61 GORFRTTAYYRGAMGIMLVYDITNKSFDNIRNIEEHSASDVERMILGKCKDL 120
QY 97 TTRKVVDTTAKFADSLGIPLETSAKNTNVEOSFMTMAEIKRKGAGTAGAERS 156
DB 121 NERQVSKRGEKLAIDYGIKLETSAKSINVEAFITLARDIMTKIKMENSLOQA 180
QY 157 NVKIOSTFVK 166
DB 181 VDLKSPVK 190

RESULT 7
US-09-925-300-1571
Sequence 1571, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1571
LENGTH: 218
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-300-1571

Query Match 45.5%; Score 403.5; DB 10; Length 218;
Best Local Similarity 39.7%; Pred. No. 1.1e-32;
Matches 85; Conservative 32; Mismatches 30; Indels 67; Gaps 4;

QY 8 DYLFKLLIGDSGVGKSCLLRFADDDYTESYSTIGVDFKIRITELDGKTIKQI--- 64
DB 24 DYLFKLLIGDSGVGKTCVLPFRSDAFNTFTISTIGDFKIRITELDGKTIKQIMDPA 83
QY 65 -----ESFNNVKOMLOEIDRYASBNVKNLLVGKCKDL 95
DB 84 AGGERFRTTAYYRGAMGIMLVYDITNKSFDNIRNIEEHSASDVERMILGKCKDL 143

[illegible]

RESULT 8

```

US-09-925-302-534
Sequence 534, Application US/09925302
Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 534
LENGTH: 246
TYPE: prt
ORGANISM: Homo sapiens
US-09-925-302-534

```

Query Match	44.1%;	Score 390.5;	DB 10;	Length 246;
Best Local Similarity	41.7%;	Pred. No. 2.5e-31;		
Matches 78; Conservative	35;	Mismatches 41;	Indels 33;	Gaps 2

```

QY      61 KLOI-----ESPNNVQWLOEIDRYASENNKL 88
      |||:
      |||:
Db      101 KIQWYTAGOERFKTITTAAYVRGAMGIIIVYDIDERSFENIQWMMKSINEMNSAGYERL 160
      |||:
      |||:
Db      41 VSAAAKAYDHLFKLLIGDSSVGKTCILIRAEADNFNNYITISTIGIDPFKRTYDIEKKI 1000
      |||:
      |||:
QY      1 MSSNNPEYFKLLIGDSSVGSCILLRADDTYESTYSTGVGFKRTIEDLCKTI 60
      |||:
      |||:
Db      41 VSAAAKAYDHLFKLLIGDSSVGKTCILIRAEADNFNNYITISTIGIDPFKRTYDIEKKI 1000
      |||:
      |||:
QY      89 LVGKKCDLITTKVYDYTTAKFEFADSLGIPELETSAKANATVGESEFMTAAEIKKRMGPGA 148
      |||:
      |||:
Db      161 LIGKKCMENAKRKQKQEOADKLAREHGIREFPETSAKSMNVDEAFSSILARDIILKSG-GR 219
      |||:
      |||:
QY      149 TAGGAER 155
      |||:
      |||:
Db      220 RSGGNGNK 226

```

RESULT 9

```

US-09-834-765-766
Sequence 766, Application US/09834765
Patent No. US20020055478A1
GENERAL INFORMATION
APPLICANT: Mary Faris
APPLICANT: Pia M. Challita-Eid
APPLICANT: Arthur B. Raitano
APPLICANT: Steve Chappell Mitchell
APPLICANT: Daniel E. H. Afar
APPLICANT: Aya Jakobovits
TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
TITLE OF INVENTION: AND DETECTION OF CANCER
FILE REFERENCE: 129.6USU1
CURRENT APPLICATION NUMBER: US/09/834,765
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: 60/197,647
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 770
SOFTWARE: fastseq for Windows Version 4.0
SEQ ID NO 766

```

```

; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-766

```

Query Match	41.98;	Score 371;	DB 10;	Length 162;
Best Local Similarity	45.18;	Pred. No. 1.2e-29;		
Matches 73;	Conservative 28;	Mismatches 29;	Indels 32;	Gaps 1;

[illegible]

RESULT 10

```

US-09/794-257-16
; Sequence 16, Application US/09794.257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; TITLE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794.257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam accession number PF00071
US-09-794-257-16

```

Query Match	40.98;	Score 362.5;	DB 10;	Length 198;
Best Local Similarity	41.7%;	Pred. No. 1.1e-28;		
Matches	86;	Conservative 25;	Mismatches 40;	Indels 55;
				Gaps 6;

[illegible]

RESULT 11

US-09-945-173-5
; Sequence 5, Application US/09945173
; Patent No. US20020127568A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Meyers, Rachel

```

: TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 38155-20035.00
: CURRENT APPLICATION NUMBER: US/09/945,173
: CURRENT FILING DATE: 2001-08-31
: PRIOR APPLICATION NUMBER: US 60/229,293
: PRIOR FILING DATE: 2000-09-01
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 198
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Consensus amino acid sequence
US-09-945-173-5

Query Match
Best Local Similarity 40.9%; Score 362.5; DB 10; Length 198;
Matches 86; Conservative 25; Mismatches 40; Indels 55; Gaps 6;

QY 13 KLLIGSGVSKSCLLRPADDTYESYISTGVDFKRTIELDKTKIKQI-----64
DB 1 KVLIGSGVSKSCLLRITDKRVEEYIPTLVDPYTKIVEGDKTKLQIMDTAGQER 60
QY 65 -----ESTNNYKQMLQEDIRYA--SEVNNKLVGNKCDLT- 97
DB 61 PVALRPAYRGAQGLLYDITSRDSFENKMKLEILRHADKQENPVLVGNKCDLED 120
QY 98 -----TKKVYDYTTAKEPADSLG-IPLETSAKNATNVEQSEFMTMAEIKRKGPG 147
DB 121 DEDLELTGQKRVVSTEGEALAKELGALPEMETSAKNTNVEAFELARILKVV--- 177
QY 148 ATAGAESKNVKTQSTPVKQSGGCC 173
DB 178 -----SEVNVNLDQ-PARKKRSKCC 196

RESULT 12
US-09-972-529-4
: Sequence 4, Application US/09972529
: Patent No. US20020150916A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: APPLICANT: Meyers, Rachel
: TITLE OF INVENTION: 47316, A NOVEL HUMAN G-PROTEIN AND USES
: TITLE OF INVENTION: THEREFOR
: FILE REFERENCE: 38155-20041.00
: CURRENT APPLICATION NUMBER: US/09/972,529
: CURRENT FILING DATE: 2002-04-08
: PRIOR APPLICATION NUMBER: US 60/237,716
: PRIOR FILING DATE: 2000-10-05
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 198
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Consensus amino acid sequence
US-09-972-529-4

Query Match
Best Local Similarity 40.9%; Score 362.5; DB 10; Length 198;
Matches 86; Conservative 25; Mismatches 40; Indels 55; Gaps 6;

QY 13 KLLIGSGVSKSCLLRPADDTYESYISTGVDFKRTIELDKTKIKQI-----64
DB 1 KVLIGSGVSKSCLLRITDKRVEEYIPTLVDPYTKIVEGDKTKLQIMDTAGQER 60
QY 65 -----ESTNNYKQMLQEDIRYA--SEVNNKLVGNKCDLT- 97
DB 61 PVALRPAYRGAQGLLYDITSRDSFENKMKLEILRHADKQENPVLVGNKCDLED 120
QY 98 -----TKKVYDYTTAKEPADSLG-IPLETSAKNATNVEQSEFMTMAEIKRKGPG 147
DB 121 DEDLELTGQKRVVSTEGEALAKELGALPEMETSAKNTNVEAFELARILKVV--- 177
QY 148 ATAGAESKNVKTQSTPVKQSGGCC 173
DB 178 -----SEVNVNLDQ-PARKKRSKCC 196

```

```

QY 98 -----TKKVYDYTTAKEPADSLG-IPLETSAKNATNVEQSEFMTMAEIKRKGPG 147
DB 121 DEDLELTGQKRVVSTEGEALAKELGALPEMETSAKNTNVEAFELARILKVV--- 177
QY 148 ATAGAESKNVKTQSTPVKQSGGCC 173
DB 178 -----SEVNVNLDQ-PARKKRSKCC 196

RESULT 13
US-09-350-874-67
: Sequence 67, Application US/09350874
: Patent No. US20020019020A1
: GENERAL INFORMATION:
: APPLICANT: Rhodes, Kenneth
: APPLICANT: An, Mengdian
: TITLE OF INVENTION: METHODS FOR TREATING CARDIOVASCULAR DISORDERS
: FILE REFERENCE: INT-069
: CURRENT APPLICATION NUMBER: US/09/350,874
: CURRENT FILING DATE: 1998-07-09
: EARLIER APPLICATION NUMBER: USSN 60/110,277
: EARLIER FILING DATE: 1998-11-30
: EARLIER APPLICATION NUMBER: USSN 60/110,033
: EARLIER FILING DATE: 1998-11-25
: EARLIER APPLICATION NUMBER: USSN 60/109,333
: EARLIER FILING DATE: 1998-11-20
: EARLIER APPLICATION NUMBER: USSN 09/298,731
: NUMBER OF SEQ ID NOS: 72
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 67
: LENGTH: 212
: TYPE: PRT
: ORGANISM: Rattus sp.
US-09-350-874-67

Query Match
Best Local Similarity 37.6%; Score 333; DB 10; Length 212;
Matches 73; Conservative 38; Mismatches 55; Indels 44; Gaps 2;

QY 8 YDLFRLKIGSGVSKSCLLRPADDTYESYISTGVDFKRTIELDKTKIKQI---64
DB 3 YALFVYIIDPTGVKSCLLDQFDRKRPVHDLTIGVEFGARMTIDKQIKQIMDT 62
QY 65 -----ESTNNYKQMLQEDIRYASPNYKLLVGNKCD 95
DB 63 AGQESRSTTSYRGAAGMLVYDITRDTFNHLLTWLEDAKQSNMAYIMLNKSD 122
QY 96 LTRKAVDYTTAKEPADSLGIPLETSAKNATNVEQSEFMTMAEIKRKGPGATAGCAEK 155
DB 123 LESRREVKKREGEAFAREHGLIPMETSAKNTNVEAFINTAKETIKQEGVFIDINNEA 182
QY 156 SNVKIQSTPV-----KQSGGCC 173
DB 183 NGIKIGPOHAATNASHGNGGQOAGGCC 212

RESULT 14
US-09-817-198A-4
: Sequence 4, Application US/09817198A
: Patent No. US20020146758A1
: GENERAL INFORMATION:
: APPLICANT: YE, Jane et al.
: TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
: FILE REFERENCE: CLO01188
: CURRENT APPLICATION NUMBER: US/09/817,198A
: CURRENT FILING DATE: 2001-03-27
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 4

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 17, 2002, 13:05:13 ; Search time 2644 Seconds
(without alignments)
1904.230 Million cell updates/sec

Title: US-09-820-003A-2

Perfect score: 886
Sequence: 1 MSSNPEDYDFKLLIGDS.....EKSNNKISTPYKSGGCC 173

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xlh
-O=/cgn2.1/USPFO.spool/US09820003/runat.13112002.134616.3176/app.query.fasta.1.327
-DB=GenEmbl -OPMT=fastap -SUPERfix=ge -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0 cdi -LIST=45
-DOCFALT=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRP=US09820003.ecgn1.1.1616.ernat.13112002.134616.3176 -NCPU=6 -ICPU=3
-NO.XLPXY -NO.MAP -LARGEOUERY -NEG SCORES=0 -MAIT -LONCLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=10 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database: GenEmbl:*

1: gb_ba:*

2: gb_hlg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_ov:*

7: gb_ov:*

8: gb_ov:*

9: gb_ov:*

10: gb_ov:*

11: gb_ov:*

12: gb_ov:*

13: gb_ov:*

14: gb_ov:*

15: gb_ov:*

16: gb_ov:*

17: gb_ov:*

18: gb_ov:*

19: gb_ov:*

20: gb_ov:*

21: gb_ov:*

22: gb_ov:*

23: gb_ov:*

24: gb_ov:*

25: gb_ov:*

26: gb_ov:*

27: gb_ov:*

28: gb_ov:*

29: em_vi:*

30: em_hlg_hum:*

31: em_hlg_inv:*

32: em_hlg_inv:*

33: em_hlg_inv:*

34: em_hlg_inv:*

35: em_hlg_inv:*

36: em_hlg_inv:*

37: em_hlg_inv:*

38: em_hlg_inv:*

39: em_hlg_inv:*

40: em_hlg_inv:*

41: em_hlg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	860	97.1	615	6 E03861	E03861 cDNA encodi
2	860	97.1	618	9 AF498929	AF498929 Homo sapi
3	860	97.1	723	9 HUWRAB1A	HUWRAB1A Homo sapi
4	860	97.1	840	4 CEFAB1	CEFAB1 Canine rab1
5	860	97.1	1428	10 MWYPT1	MWYPT1 Mouse mRN
6	860	97.1	1433	9 BC000905	BC000905 Homo sapi
7	860	97.1	1444	10 BC002077	BC002077 Mus muscu
8	860	97.1	2686	10 AF226873	AF226873 Mus muscu
9	849	95.8	618	10 RATRASA	RATRASA Rat ras-rel
10	838	94.6	841	5 DYGORAB1	DYGORAB1 Sequence
11	768.5	86.7	925	6 AR070365	AR070365 Sequence
12	768.5	86.7	1985	6 AX086155	AX086155 Sequence
13	768.5	86.7	1985	9 HSM801608	HSM801608 Homo sapi
14	762.5	86.1	654	10 RMRAB1B	RMRAB1B Homo sapi
15	762.5	86.1	161955	2 AC106368	AC106368 Rat cDNA fo
16	761.5	85.9	1859	10 BC016408	BC016408 Mus muscu
17	758.5	85.6	131988	2 AC123143	AC123143 Rat cDNA no
18	740.5	83.6	1988	2 HSM245875	HSM245875 Homo sapi
19	737.5	82.8	185285	2 AC094411	AC094411 Homo sapi
20	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
21	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
22	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
23	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
24	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
25	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
26	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
27	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
28	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
29	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
30	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
31	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
32	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
33	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
34	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
35	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
36	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
37	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
38	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
39	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
40	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi
41	733.5	82.8	185285	2 AC139414	AC139414 Homo sapi

RESULT 1

ALIGNMENTS

E03861
LOCUS E03861 615 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding GTP binding protein ypt1.
ACCESSION E03861
VERSION E03861.1 GI:2172075
KEYWORDS JP 1992226000-A/3.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
REFERENCE 1 (bases 1 to 615)
AUTHORS Yamaguchi,K., Uchida,K., Takai,Y., Aisaka,K. and Sakurada,K.
TITLE LOW-MOLECULAR WEIGHT PROTEIN BINDING TO GUANOSINE TRIPHOSPHATE
JOURNAL Patent: JP 1992226000-A 3 14-AUG-1992;
KYOMA HAKKO KOGYO CO LTD
OS Mus sp. (mouse)
PN JP 1992226000-A/3
PD 14-AUG-1992
PF 25-APR-1991 JP 1991095304
PR 21-MAY-1990 JP 90P 130547
PI YAMAGUCHI KAZUO, UCHIDA KAZUHISA, TAKAI YOSHIMI, AISAKA KAZUO,
PI SAKURADA KAZUHIRO
PC C07K15/18,C07K13/00,C07K15/06,C12N15/12,G01N33/50//A61K37/02;
CC strandness: Single;
topology: Linear;
location/Qualifiers
1..615
/organism="Mus sp."
/db_xref="taxon:10095"
BASE COUNT 196 a 117 c 153 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 4.51e-75 Length: 615
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best Local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
Gaps: 1

US-09-820-003a-2 (1-173) x E03861 (1-615)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
|||||
Db 1 ATGTCAGCATGAAATCCCGAATATGATTAATTCAGTACTTGATGGCGATTCT 60
QY 21 GlyValGlyLysSerCysLeuLeuArgPheAlaAspThrTyrThrGluSerTyr 40
|||||
Db 61 GGGGTGGAAAGTCCCTGCTCTCTTAAGTTGGAGATGATACGTATACGAAAGCTAC 120
QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
|||||
Db 121 ATGAGCAAAATGGTGATGATTTCAGATACGAACTATAGATTAGATGGGAAACAAATC 180
QY 61 LysLeuGlnIle----- 64
|||||
Db 181 AAGCTAGATATGAGACACAGACAGCCAGAAAGATTTCGAACAATCACTTCAGATTAT 240
QY 65 -----GluSerPheasn 68
|||||
Db 241 TACAGAGAGCCCATGCGATCATAGTGTATGATGTGACAGATACAGAGTCCCTCAAT 300
QY 69 AsnValLysGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
|||||
Db 301 AACGTTAAACAGTGGCTCCAGAGATAGATCGCTACGCCAATGAATAATGTCAACAAGTTG 360
QY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108
|||||
Db 361 TTGGTAGGAGCAAAATGTGACCTGCACCAAAAGAAAGTAGAGTACACAAAGCAAG 420
QY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluTyrSerAlaLysAsnAlaThrAsn 128
|||||
Db 421 GAATTTGGAGATCCCTTGGAATTCATTTTGGAAACCACTGCTAAGAACGCAACGAAAT 480

QY 129 ValGlnSerPheMetThrMetAlaGluIleLysLysArgMetGlyProGlyAla 148
|||||
Db 481 GTAGACAGCTTTTCATCAGATGCGAGCTAGATTAAAGCAATGGGCTCCGAGCT 540
QY 149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
|||||
Db 541 ACACCTGCTGCTGCCGAGAAAGTCCCAATGTTAAATCCAGAGCACTCCAGTCACAGCTCA 600
QY 169 GlyGlyGlyCysCys 173
|||||
Db 601 GGTGAGGCTCTCTCC 615

RESULT 2
AF498929
LOCUS AF498929
DEFINITION 618 bp mRNA linear PRI 01-MAY-2002
ACCESSION AF498929
VERSION AF498929.1 GI:20379033
KEYWORDS cds.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 618)
AUTHORS Puhl,H.L. Iii, Ikeda,S.R. and Aronstam,R.S.
TITLE Homo sapiens RAB family small GTP binding protein RAB1A
JOURNAL Unpublished
2 (bases 1 to 618)
AUTHORS Puhl,H.L. Iii, Ikeda,S.R. and Aronstam,R.S.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) CDNA Resource Center, Guthrie Research
Institute, One Guthrie Square, Sayre, PA 18840, USA
location/Qualifiers
1..618
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
1..618
/gene="RAB1A"
1..618
/gene="RAB1A"
/note="RAB family member"
/codon_start=1
/product="small GTP binding protein RAB1A"
/protein_id="P46210.7.1"
/db_xref="GI:20379034"
/translation="MSSMNPEDYDLKLLIGSGVSKSCLLRFADDTYESISTI
GVDFKIRTIIDDKTKIKLIQIMDPAQGERFRTISYYRGAHGIIVYDVYDQSPFNV
KWLQIEDRYASENVNKLIVGNCGLDTTKKIVDYTTAKERADSLGIPIFLETSKRNATN
VEQSPFMMAEIKRMGDPGATAGAAESNVKIGSTPKVGGGCC"

BASE COUNT 203 a 110 c 149 g 156 t
ORIGIN

Alignment Scores:
Pred. No.: 4.53e-75 Length: 618
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best Local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
Gaps: 1

US-09-820-003a-2 (1-173) x AF498929 (1-618)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
|||||
Db 1 ATGTCAGCATGAAATCCCGAATATGATTAATTCAGTACTTGATGGCGACTCA 60
QY 21 GlyValGlyLysSerCysLeuLeuArgPheAlaAspThrTyrThrGluSerTyr 40
|||||
Db 61 GGGGTGGAAAGTCTGCTCTCTTAAGTTGGAGATGATACGTATACGAAAGCTAC 120

```

OY 41 IleserThriIleGIyValAspPheLysIleArgThrIleGIuLeuAspGlyLysThrIle 60
Db 121 ATCACACAAATTCGTGTGATTTCAAAATAGACTATAGGTTAGACGGGAACAATC 180
OY 61 LysLeuGlnIle----- 64
Db 181 AACGCTCAATATGGGACACAGACGCCAGGAAGAAATTCACATCACCCTCAGTTAT 240
OY 65 -----GluSerPheAsn 68
Db 241 TACAGAGAGCCCATGCAATCATGTTGTTATGATGACAGATCAGAGAGCTTCAAT 300
OY 69 AsnValLysGlnTrpLeuGlnGluIleAspArgTrpAlaSerGluAsnValAsnLysLeu 88
Db 301 AATGTTAAACAGGGCTGCGAGAAATAGATCGTTATGCAAGTGAATTCACAACAAATG 360
OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysValValAspTrpThrThrAlaLys 108
Db 361 TTGGTAGGGAACAATGATGATCTACCAACAAAGATAGTACACACAGCAGAG 420
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluTrpSerAlaLysAsnAlaThrAsn 128
Db 421 GAATTTGCTGATTCCTTGGAATTCCTTTTGGAACACAGTGTGAATGCAACGAAAT 480
OY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
Db 481 GTAGACACAGCTTTCATGACATGCGAGTGAATTAAGCAATGGTCCCGAGCA 540
OY 149 ThrAlaGlyLysAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
Db 541 ACAGCTGGTGTGCTGAGACATGCAATGTTAAATTCAGACAGCTCAGTCAAGAGTCA 600
OY 169 GlyGlyGlyCysCys 173
Db 601 GGTGAGGTGCTGC 615

RESULT 3
HUMRABIA 723 bp mRNA linear PRI 08-JAN-1995
LOCUS Homo sapiens GTP-binding protein (RAB1) mRNA, complete cds.
DEFINITION M28209 J04941
VERSION M28209.1 GI:550059
KEYWORDS GTP-binding protein; ras oncogene.
SOURCE Homo sapiens (tissue library: of J.Mallet) pheochromocytoma CDNA to
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 723)
AUTHORS Zahraoui, A., Touchot, N., Chardin, P. and Tavilian, A.
TITLE The human Rab genes encode a family of GTP-binding proteins related
JOURNAL to yeast Yrp1 and Sec4 products involved in secretion
MEDLINE J Biol Chem. 264 (21), 12394-12401 (1989)
PUBMED 8930866
COMMENT 2501306
FEATURES
SOURCE On Sep 27, 1994 this sequence version replaced gi:341862.
Location/Qualifiers
1..723
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Unassigned"
/tissue_type="pheochromocytoma"
/tissue_lib="of J.Mallet"
1..723
/gene="RAB1"
51..668
/gene="RAB1"
/codon_start=1
/product="GTP-binding protein"
/protein_id="AA060240.1"
/db_xref="GI:550060"
/db_xref="GDB:G00-118-857"
/translation="MSSMNPEDYILFKLLIGDSVGKSCLLRFADDTYESTYSTI

```

```

GUDFKRTIEIDKTYIKLOHDTAQGEFETTSYRGAGITIVYVDYDQSENNV
KQKLGIDIRYASERVNKLIVGNKCDLITKRYVDITTAKEFASGLIFPLETSKNTIN
VDSFNTMAELTKRMGPATAGAGAKSNVATQSTIVKQSGGCC"
BASE COUNT 225 a 142 c 180 g 176 t
ORIGIN
Alignment Scores:
Pred. No.: 5,46e-75 Length: 723
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best Local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
DB: Gaps: 1
US-09-820-003a-2 (1-173) x HUMRABIA (1-723)
OY 1 MetSerSerMetAsnProGluTrpAspTrpLeuPheLysLeuLeuIleGlyAspSer 20
Db 51 ATGTCAGCATGAATCCCGAATATATATTATTCAGTTCCTTGATGGCCGACTCA 110
OY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTrpTrpGluSerTyr 40
Db 111 GCGCTTGGAAGACTTCTCCCTCTCTTCTAGGTTGCAGATGATATATACGAAAGCTAC 170
OY 41 IleserThriIleGIyValAspPheLysIleArgThrIleGIuLeuAspGlyLysThrIle 60
Db 171 ATCAGCACAAATGCTGTGATTTCAAAATAGAACTATAGAGTTAGACGGGAACAATC 230
OY 61 LysLeuGlnIle----- 64
Db 231 AACGCTCAATATATGGACACAGCAGCAGGAAAGATTCGACATCACCCTCAGTTAT 290
OY 65 -----GluSerPheAsn 68
Db 291 TACAGAGAGCCCATGCAATCATGTTGTTATGATGTGACAGATCAGAGAGTCTTCAT 350
OY 69 AsnValLysGlnTrpLeuGlnGluIleAspArgTrpAlaSerGluAsnValAsnLysLeu 88
Db 351 AATGTTAAACAGTGGCTGCGAGAAATAGATCGTTATGCCAGTGAATGTCACAAATG 410
OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysValValAspTrpThrThrAlaLys 108
Db 411 TTGCTAGGGAACAATGATGATCTGACCAACAAAGATGATGACTACACACAGCAGAG 470
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluTrpSerAlaLysAsnAlaThrAsn 128
Db 471 GAATTTGCTGATTCCTTGGAATTCCTTTTGGAACACAGTGTGAATGCAACGCAAC 530
OY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
Db 531 GTAGACACAGCTTTCATGACATGCGAGTGAATTAAGCAATGGTCCCGAGCA 590
OY 149 ThrAlaGlyLysAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
Db 591 ACAGCTGGTGTGCTGAGACATGCAATGTTAAATTCAGACAGCTCAGTCAAGAGTCA 650
OY 169 GlyGlyGlyCysCys 173
Db 651 GGTGAGGTGCTGC 665

RESULT 4
LOCUS CERAB1 840 bp mRNA linear MAM 30-SEP-1999
DEFINITION Canine rab1 mRNA for ras-related GTP-binding protein involved in
X56384 membrane traffic.
VERSION X56384.1 GI:913
KEYWORDS GTP-binding protein; protein transport; ras gene; ras gene family.
SOURCE Canis familiaris.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 840)

```

AUTHORS Chavrier, P., Vingron, M., Sauder, C., Simons, K. and Zerial, M.
 TITLE Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line
 JOURNAL Mol. Cell. Biol. 10 (12), 6578-6585 (1990)
 MEDLINE 91061765
 PUBMED 2123294
 REFERENCE 2 (pages 1 to 840)
 AUTHORS Zerial, M.
 TITLE Direct Submission
 JOURNAL Submitted (17-AUG-1990) Zerial M., EMBL, Meyerhofstrasse 1, 6900 Heidelberg, FRG

FEATURES
 source
 location/Qualifiers
 1..840
 /organism="Canis familiaris"
 /strain="cocker spaniel"
 /db_xref="taxon:9615"
 /clone="47"
 /cell_line="MDCK"
 149..766
 /gene="rab1"
 149..766
 /gene="rab1"
 /function="GTP binding protein"
 /codon_start=1
 /protein_id="CAB56775.1"
 /db_xref="GI:6006435"
 /translation="MSSMNPEDYDLFKLLIGDSGVKSCLLRFADDTYESIST
 GVDFKRTIEDLGGTKTIKLOIWDPAQGERFTTSSYRGHGIIVYDWDQSFNNV
 KQRLQETDRASENVKLLVGNKCDLTTKKVYVDTTAKERADSLGIPLETSAKNNATN
 VEOSFMTMAEIKRRMPCGATAGAGAEKSNVKIOSTVPKOSGGCC"

BASE COUNT 250 a 176 c 220 g 194 t
 ORIGIN

Alignment Scores:
 Pred. No.: 6.52e-75 Length: 840
 Score: 860.00 Matches: 173
 Percent Similarity: 84.39% Conservative: 0
 Best Local Similarity: 84.39% Mismatches: 0
 Query Match: 97.07% Indels: 32
 DB: 4 Gaps: 1

US-09-820-003a-2 (1-173) x CFRAB1 (1-840)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20
 |||||
 Db 149 ATGTCAGAGCATGATCCGAGATATGATTATTATCAAGTACTTCGATGGCGACT 208
 |||||
 QY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40
 |||||
 Db 209 GGGCTTGGAAGTCTGCTCTCTTAGGTTTCAGATGATACATATACAGAAAGCTAC 268
 |||||
 QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGlyLeuAspGlyLysThrIle 60
 |||||
 Db 269 ATCGGCCCAATGGTGTGGATTTCAAAATAGAACTATAGATTAGATGGGAAAAACAATC 328
 |||||
 QY 61 LysLeuGlnIle----- 64
 |||||
 Db 329 AAGCTTCAAAATATGGACACAGACAGGCCCAAGAAATTTGCAACATCACTCCAGTTAT 388
 |||||
 QY 65 -----GluSerPheAsn 68
 |||||
 Db 389 TACAGAGAGCCCATGCAATCATAGTGTGTATGACGTGACAGATTCAGAGTCCCTTCAT 448
 |||||
 QY 69 AsnValLysGlnTyrPheGlnGlnIleAspArgTyrIleAspArgIleValAsnValAsnLysLeu 88
 |||||
 Db 449 AATGTTAAACGTGGTGGCGAGAAATAGACCGTTATGCGCATGAAACGTTCAACAGCTTG 508
 |||||
 QY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108
 |||||
 Db 509 TTGGTAGGAAACAATGCGATCTGACCAACAAGAAAGTAGAGACTACACACAGCAAG 568
 |||||
 QY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
 |||||

Db 569 GAATTTGGCGGATTCCTTGGAATTCATTTTGGAAACCACTGCTTAAGACCAACAAT 628
 QY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
 |||||
 Db 629 GTAGACAGCTTTTCATGACAGATGCGACGTACAGTTAAAAAGCAATGGGTCTGGAGCA 688
 |||||
 QY 149 ThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
 |||||
 Db 689 ACACGTGGTGGTCAGAGAAAGTCCAAATGTAATTCAGACACTCCGCTCAAGACATCA 748
 |||||
 QY 169 GlyGlyGlyCysCys 173
 |||||
 Db 749 GTTGAGAGTGTGCTGC 763

RESULT 5
 MMYPT1 1428 bp mRNA linear ROD 12-SEP-1993
 LOCUS MMYPT1
 DEFINITION Mouse mRNA for ras-related YPT1 protein.
 ACCESSION Y00094
 VERSION Y00094.1 GI:55456
 KEYWORDS ras-related protein; YPT1 gene.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1428)
 REFERENCES
 AUTHORS Gallwitz, D.
 TITLE Direct Submission
 JOURNAL Submitted (13-NOV-1987) Gallwitz D., Max-Planck-Institut fuer Biophysikalische Chemie, Abt. Molekulare Genetik, Postfach 2841, D-3400 Goettingen, FRG
 2 (bases 1 to 1428)
 Haubruck, H., Disela, C., Wagner, P. and Gallwitz, D.
 TITLE The ras-related ypt1 protein is an ubiquitous eukaryotic protein: isolation and sequence analysis of mouse cDNA clones highly homologous to the yeast YPT1 gene
 JOURNAL EMBO J. 6 (1987) in press
 FEATURES
 source
 location/Qualifiers
 1..1428
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_line="F9 and C3H10T1/2"
 130..747
 /note="YPT1 protein (AA 1-205)"
 /codon_start=1
 /protein_id="CAA68284.1"
 /db_xref="GI:55457"
 /db_xref="MSD:W61:97842"
 /db_xref="SWISS-PROT:P11476"
 /translation="MSSMNPEDYDLFKLLIGDSGVKSCLLRFADDTYESISTI
 GVDFKRTIEDLGGTKTIKLOIWDPAQGERFTTSSYRGHGIIVYDWDQSFNNV
 KQRLQETDRASENVKLLVGNKCDLTTKKVYVDTTAKERADSLGIPLETSAKNNATN
 VEOSFMTMAEIKRRMPCGATAGAGAEKSNVKIOSTVPKOSGGCC"
 1401..1406
 /note="pol. polyA signal"
 1428
 /note="pol. polyA site"
 polyA_site
 BASE COUNT 410 a 263 c 324 g 431 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.22e-74 Length: 1428
 Score: 860.00 Matches: 173
 Percent Similarity: 84.39% Conservative: 0
 Best Local Similarity: 84.39% Mismatches: 0
 Query Match: 97.07% Indels: 32
 DB: 10 Gaps: 1

US-09-820-003a-2 (1-173) x MMYPT1 (1-1428)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20
 |||||
 Db 130 ATGTCAGAGCATGATCCGAGATATGATTATTATCAAGTACTTCGATGGCGACTTCT 189

|||||
Db 659 GTAAACAGCTTTTCATGACGATGCGACCTGAGATTAAAAAGCCAAATGGGCGCCGAGACA 718
Oy 149 ThrAAGlGyAlAGlulYsSerAsnValYsIleGInSerThrProValYsGIser 168
Db 719 ACACCTGCTGCTGAGAAAGTCCAAATGTTAAATTCAGACACTCCAGTCAAGCAGTCA 778
Oy 169 GYglYglYcYcYs 173
Db 779 GGTGAGGTTCTCTGC 793
RESULT 7
BC002077
LOCUS BC002077 1444 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, RAB1, member RAS oncogene family, clone MGC:6226
IMAGE:3592802, mRNA, complete cds.
ACCESSION BC002077
VERSION BC002077.1 GI:12805232
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1444)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
REMARK Contact: MGC help desk
COMMENT Email: cgapsb-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 11 Row: a Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6679586.
FEATURES
Source
1. 1444
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:6226 IMAGE:3592802"
/tissue_type="Mammary tumor, Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI_CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
196..813
CDS
/codon_start=1
/product="RAB1, member RAS oncogene family"
/protein_id="AAH02077.1"
/db_xref="GI:12805233"
/db_xref="locusid:5861"
/translation="MSSMNPEDYLFKLLIGSGVKSCLLRPADDTYESYIST
GVDPKIRITLDGKTKLQIWDYAGQERFTTISYRGHGLIVYDVTDSEFNW
KWLQIEDRYASENVKLLVGNKCDLTTKKVVDYTTAKERADSLGIPLETSAKNATN
VEOSFTMAEIKRMGPGATAGAEKSNVKIOSTPVKOSGGCC"
BASE COUNT 412 a 270 c 342 g 420 t
ORIGIN
Alignment Scores:
4 47

Pred. No.: 1,24e-74 Length: 1444
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
DB: 10 Gaps: 1
US-09-820-003a-2 (1-173) x BC002077 (1-1444)
Oy 1 MetSerSerMetAnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
Db 196 ATGTCCAGCATGAATCCCGAATATGATTATTATTCAGTTACTTCTGATGGCGATTCT 255
Oy 21 GYAlAGlYsSerCYsLeuLeuLeuArPheAlaAspAspThrTYrThrGluSerTYr 40
Db 256 GGGTGTGGAAGTCTGCTGCTCTCTTACGTTTCAGATGATGATAGCAAGAAAGCTAC 315
Oy 41 ILeSerThrIleGlyAlaAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
Db 316 ATCAACCAATGGTGTGGATTTCAGATACGAACATATGAGTTAGATGGGAAAAACATC 375
Oy 61 LysLeuGlnIle-----
Db 376 AAGCTACAGATATATGGACACACAGCCAGCAAGAAATTTCGAACATTCACCTTCAGTTAT 435
Oy 65 -----GluSerPheAsn 68
Db 436 TACAGAGAGCCCATGCGATCATGATGTTGTATGATGATGACAGATCAGAGATCCTTCAT 495
Oy 69 AsnValYsGlnTTPLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
Db 496 AACGTTAAACAGTGGCGCAGAGATGATCGTACGCCAGTGAATAATGTCACAAAGTTG 555
Oy 89 LeuValGlyAsnLysCYsAspLeuThrTrpLysValValaAspTyrThrThralaLys 108
Db 556 TTGGTAGGAACAATGTGACCTACCCACAAGAAAGTAGAGCTACACACAGCAAG 615
Oy 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
Db 616 GAATTCGACATGCCCTTGGAATTCATTTTGGAAACCAAGTGTAAAGCAACCAAGAT 675
Oy 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
Db 676 GTAAACAGCTTTTCATGACGATGCGACGTCGAGATTAAAAAGCAAGGGCTCGGAGCT 735
Oy 149 ThrAAGlGyAlAGlulYsSerAsnValYsIleGInSerThrProValYsGIser 168
Db 736 ACACCTGCTGCTGAGAAAGTCCAAATGTTAAATTCAGACACTCCAGTCAAGCAGTCA 795
Oy 169 GYglYglYcYcYs 173
Db 796 GGTGAGGTTCTCTGC 810
RESULT 8
AF226873
LOCUS AF226873 2686 bp mRNA linear ROD 13-FEB-2000
DEFINITION Mus musculus small GTP-binding protein RAB1A mRNA, complete cds.
ACCESSION AF226873
VERSION AF226873.1 GI:6969621
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2686)
AUTHORS Wu, G. and Dorn, G.W. II.
TITLE Mouse RAB1A, member of RAS oncogene family mRNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2686)
AUTHORS Wu, G. and Dorn, G.W. II.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-2000) Internal Medicine, University of
Cincinnati, 231 Bethesda Avenue, Cincinnati, OH 45267-0590, USA

FEATURES		Location/Qualifiers	
source	1..2686	/organism="Mus musculus"	
	/db_xref="taxon:10090"		
CDS	29..646	/note="member of RAS oncogene family"	
	/codon_start=1		
	/product="small GTP-binding protein RAB1A"		
	/protein_id="AA03844.1"		
	/db_xref="GI:6969622"		
	/translation="MSSMNPEDYFLKLLIDSGVSKCLLRFADDTYESYISTIGVDFKIRIPLDGGTKIKLQIMDTAGOEFRITSSYRGAGIIVYDITDSEFNWKMQLDIRASENVNKLIVGNKCDLTKRVVDYTTAEKPADSLGIFLETSAKNATNVEQSPMTMAEIKRMKGAVGAGAEKSNVKIQTSPVKSQSGGCC"		
BASE COUNT	778 a 520 c 563 g 825 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.58e-74	Length:	2686
Score:	860.00	Matches:	173
Percent Similarity:	84.39%	Conservative:	0
Best Local Similarity:	84.39%	Mismatches:	0
Query Match:	97.07%	Indels:	32
DB:	10	Gaps:	1
US-09-820-003a-2 (1-173) x AF226873 (1-2686)			
OY	1	MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer	20
DB	29	ATGTCACGATGATCCGAAATGATATTATTCACAGTTACTTCTGATGGCGATTC	88
OY	21	GlyValIleGlySerCysLeuLeuLeuArgPheAlaAspThrTyrThrGluSerTyr	40
DB	89	GGGTTGGAAAGCTCCCTCTCTAGCTTTCAGATGATGATACGAAAGCTAC	148
OY	41	IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle	60
DB	149	ATCAGCACAATGGTGCTGATTCAGATACGACATACAGTAAAGTAAAGCAACATC	208
OY	61	LysLeuGlnIle-----	64
DB	209	AAGCTACAGATATGGGACACAGCAGCCAGAAAGATTTCGAACATCATCTCCAGTTA	268
OY	65	-----GluSerPheAsn	68
DB	269	TACAGAGAGCCCATGATCATAGTTGTATGATGTGACAGATCAGAGAGTCTTCATAT	328
OY	69	AsnValLysGlnTyrPleuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu	88
DB	329	AACGTTAAACAGGCTCCAGAGATAGATGCTACGCCAGTAAATGTCCACAACTTG	388
OY	89	LeuValIleLysAsnLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLys	108
DB	389	TTGGTAGGGAACAAATGTGACCTGACCAAGAAAGTAGTACATACACAAAGCAAAAG	448
OY	109	GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn	128
DB	449	GAATTTCCAAATTCCTTGGAATTCATTTTGGAAACAGTCTTAAGAACGCAACCAAT	508
OY	129	ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyVala	148
DB	509	GTAACAACAGCTTTATCATCAGATGACGCTGAGATTAAGGAAATGGGCTCGGAGCT	568
OY	149	ThrAlaGlyValAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer	168
DB	569	ACACCTGTGGTGCCGAGAAATCCAAATGTTAAATCCAGAGACCTCAGTCAACGACTCA	628
OY	169	GlyLysGlyCysCys 173	
DB	629	GGTGAAGGCTGCTGC 643	
RESULT 9			
RATRASA			
LOCUS			
DEFINITION	RATRASA	618 bp	mrna
ACCESSION	Rat ras-related protein mRNA, clone NTRAB1R.		
VERSION	J02998.1		
KEYWORDS	ras oncogene.		
SOURCE	Rat brain, cDNA to mRNA, (library of D. Anderson), clone NTRAB1R.		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 618)		
AUTHORS	Touchoot, N., Chardin, P. and Tavillan, A.		
TITLE	Four additional members of the ras gene superfamily isolated by an oligonucleotide strategy: molecular cloning of vpr-related cDNAs from a rat brain library		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 84 (23), 8210-8214 (1987)		
MEDLINE	88068563		
PUBMED	3317403		
COMMENT	Draft entry and computer-readable sequence for [1] kindly provided by N.Touchoot, 01-OCT-1987.		
FEATURES	Location/Qualifiers		
source	1..618		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	1..618		
	/note="ras protein"		
	/codon_start=1		
	/protein_id="AAA42006.1"		
	/db_xref="GI:206553"		
	/translation="MSSMNPEDYFLKLLIDSGVSKCLLRFADDTYESYISTIGVDFKIRIPLDGGTKIKLQIMDTAGOEFRITSSYRGAGIIVYDITDSEFNWKMQLDIRASENVNKLIVGNKCDLTKRVVDYTTAEKPADSLGIFLETSAKNATNVEQSPMTMAEIKRMKGAVGAGAEKSNVKIQTSPVKSQSGGCC"		
BASE COUNT	194 a 124 c 160 g 140 t		
ORIGIN	unreported.		
Alignment Scores:			
Pred. No.:	5.34e-74	Length:	618
Score:	849.00	Matches:	171
Percent Similarity:	83.41%	Conservative:	0
Best Local Similarity:	83.41%	Mismatches:	2
Query Match:	95.82%	Indels:	32
DB:	10	Gaps:	1
US-09-820-003a-2 (1-173) x RATRASA (1-618)			
OY	1	MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer	20
DB	1	ATGTCACGATGATCCGAAATGATATTATTCACAGTTACTTCTGATGGCGACTCT	60
OY	21	GlyValIleGlySerCysLeuLeuLeuArgPheAlaAspThrTyrThrGluSerTyr	40
DB	61	GGGTTGGAAAGCTCCCTCTCTAGCTTTCAGATGATGATACGAAAGCTAC	120
OY	41	IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle	60
DB	121	ATCAGCACAATGGTGCTGATTCAGATACGACATACAGTAAAGTAAAGCAACATC	180
OY	61	LysLeuGlnIle-----	64
DB	181	AAGCTTACAGATATGGGACACAGCAGCCAGAAAGTTTCGAACATCATCTCCAGTTA	240
OY	65	-----GluSerPheAsn	68
DB	241	TACAGAGAGCCCATGATCATAGTTGTATGATGTGACAGCACGAGAGTCTTCATAT	300
OY	69	AsnValLysGlnTyrPleuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu	88
DB	301	AACGTTAAACAGTGGCTCCAGAGATAGATGCTACGCCAGTAAATGTCCAAACAGTTG	360
OY	89	LeuValIleLysAsnLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLys	108
DB	361	TTGGTAGGGAACAAATGTGACCTGACCAAGAAAGTAGTAGTACACAAAGCCACAG	420

QY	109	GIUPEHLEAASPSERLENGLYLTERPHOLENGUJUTHTNSERALALYASALATHRSN	128
Db	421	CAATTTCAGATGATCCCTTGGAATTCATTTTGGGAACCAAGCTCAAGACAAAGAT	480
QY	129	ValGluGlnSerPheMetThrMetAlaAlaGluIleTyrSLysArgMetGlyProGlyAla	148
Db	481	GTAGAACAGTCTTTCATGACCATGATGCACGCGAGATTTAAAGCGGATGGTCTTGAGCA	540
QY	149	ThrAlaGlyGlyAlaGluLysSerSerValLysIleGlnSerThrProValLysGlnSer	168
Db	541	ACAGCTGAGGCGGAAAGTCATGTTAAATTCAGAGCATCTCAGTCAAGCATCA	600
QY	169	GIYGLIYGLYCYSQYS	173
Db	601	GGTGGAGCTCTCTGC	615
RESULT	10		
DYGORABI			
LOCUS		841 bp	mRNA linear VRT 28-Apr-1993
DEFINITION		Discopygge ommata GTP-binding protein (o-rab1) mRNA, complete cds.	
ACCESSION		M38393	
VERSION		M38393.1	GI:213122
KEYWORDS		GTP-binding protein.	
SOURCE		D.ommata electric lobe electromotor neuron, cDNA to mRNA.	
ORGANISM		Discopygge ommata	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Squalae; Hymnosqualeae; Pristiogastera; Batoidae; Torpediniformes; Narcinoidae; Narcinidae; Discopyge.	
ADDITIONAL		1 (bases 1 to 841)	
TITLE		Ngsee,J.K., Elferink,L.A. and Scheller,R.H.	
JOURNAL		A family of ras-like GTP-binding proteins expressed in electromotor neurons	
COMMENT		Unpublished (1990)	
		Draft entry and computer-readable sequence for [1] kindly submitted by J.K.Ngsee, 10-SEP-1990.	
		Herrin Lab	
		Stanford University	
		Stanford, CA 94305-5020	
		Draft entry and computer-readable sequence for [1] kindly submitted by J.K.Ngsee, 10-SEP-1990.	
		Dept of Biological Sciences	
		Herrin Lab	
		Stanford University	
		Stanford, CA 94305-5020.	
FEATURES		Location/Qualifiers	
source		1..841	
		/organism="Discopyge ommata"	
		/db_xref="taxon:7785"	
		/cell_type="electromotor neuron"	
		/tissue_type="electric lobe"	
		1..841	
gene		/gene="o-rab1"	
		79..687	
CDS		/gene="o-rab1"	
		/codon_start=1	
		/product="GTP-binding protein"	
		/protein_id="AAA49234.1"	
		/db_xref="GI:213123"	
		/translation="MNPEDYDLFKLLIGDSGVKSCLLLEFADDTYTESYISTIGVD	
		FKRTIELDGETTKLQIWDPAQGERFPTITSSYRGAAHGIIVYDVLDQSEFNVRKQM	
		LQIEDRYASENVNKLIVGNKCDLITKRVVDTAKERFADSLGIPFETSAKNATNVQD	
		AFMTMAEIKRMRPGATNGSGSEKSNVINIDSTPVKSGSGGCG"	
BASE COUNT	269 a	154 c	194 g
ORIGIN			224 t
Alignment Scores:	9.05e-73	Length:	841
Pred. NO.:	838.00	Matches:	167
Score:	83.41%	Conservative:	4
Percent Similarity:	81.46%	Mismatches:	2
Best Local Similarity:	94.58%	Indels:	32

DB:	5	Gaps:	1
US-09-820-003A-2 (1-173) x DYCORABI (1-841)			
Qy 1 MetSerSerMetAsnProGluuTyraSPtyrLeuPheLysLeuLeuLeuIleGlyAspSer 20			
Db 70 ATGTCAACCATGAATCCCGAATATGACATATTATTCACAGCTGCTGTGATTGGACATTCA 129			
Qy 21 GlyValIGlyLysSerCysLeuLeuLeuArgPheAlaAspSerThrTyrThrGluSerTyr 40			
Db 130 GGTGTGGAAAAATCTGTCTCTCTCTGTGATTTCAGATGATACATATACAGAAAGTTAT 189			
Qy 41 IleserTrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60			
Db 190 ATCACTACCAATTTGGTGTGATTATAAATACAGAACAAATAGACTTAGACGGCAGACCATC 249			
Qy 61 LysLeuGlnIle----- 64			
Db 250 AAACCTCAAACTCTGGGACACGCGCTGTGACGAGCGGTTTCGAACAATCCAGCTCCAGTTAC 309			
Qy 65 -----GluSerPheAsn 68			
Db 310 TACAGAGGGGACATGCATCATAGTTGTGTATGATGTACAGACAGGAGTCCGTTTAA 369			
Qy 69 AsnValLysGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88			
Db 370 AATGTAACCAATGAGCTTCAGGAAATAGATCGTATGCGAGGAAATGTTAAACAAGTTA 429			
Qy 89 LeuValIGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThraLalys 108			
Db 430 CTGGTGGGAACAATGATGATCTGACACCAACAAAGAGTGTGATTATACACACCAAG 489			
Qy 109 GluPheAlaAspSerLeuGlyTLeProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128			
Db 490 GAATTTCGACAGCTCCCTTGGGATCCCATTTTGGAAACAAAGTCAAACATCCAAAT 549			
Qy 129 ValGlnGlnSerPheMetThrMetAlaIleGluIleLysLysArgMetGlyProGlyAla 148			
Db 550 GTAGAACCAAGCTTCATGACTATGAGCTGCTGAGATTAAAAACGAATGGGTCTGGAGCT 609			
Qy 149 ThrAlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168			
Db 610 ACATCTGGTGGTTCGAAAGTCAATGTAAATATCCAAAGCAGCCAGTGAAGTCTGCT 669			
Qy 169 GlyGlyGlyCysCys 173			
Db 670 GGTGGGGCTGCTGCTG 684			
RESULT 11			
LOCUS AR070365 925 bp DNA linear PAT 18-FEB-2000			
DEFINITION Sequence 4 from patent US 5892012.			
ACCESSION AR070365			
VERSION AR070365.1 GI:7221253			
KEYWORDS			
SOURCE Unknown.			
ORGANISM Unknown.			
REFERENCE 1 (bases 1 to 925)			
AUTHORS Hillman,J.L., Lal,P., Corley,N.C. and Shah,P.			
TITLE Rab Proteins			
JOURNAL Patent: US 5892012-A 4 06-APR-1999;			
FEATURES location/Qualifiers			
source 1..925			
BASE COUNT 198 a 263 c 288 g 176 t			
ORIGIN /organism="unknown"			
Alignment Scores:			
Pred. No.: 5,92e-66			925
Score: 768.50			Matches: 156
Percent Similarity: 80.39%			Conservative: 8
Best Local Similarity: 76.47%			Mismatches: 7

[illegible]

Pred. No.:	1,46e-65	Length:	1985
Score:	768.50	Matches:	156
Percent Similarity:	80.39%	Conservative:	7
Best Local Similarity:	76.47%	Mismatches:	8
Query Match:	86.74%	Indels:	33
DB:	6	Gaps:	2
US-09-820-003A-2 (1-173) x AX086155 (1-1985)			
QY	2	SeSerMetcAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGly	21
DB	42	GGCGGCATGGAACCCGGAATATGATACCTGTTTAAGCTCTTTGATATGGCGATCAGCG	101
QY	22	ValGlySerCysLeuLeuLeuAtrPheAlaAspSerPthrTyrTrhgusertYrle	41
DB	102	GTGGGCAACGATCGCTCTCTCGCGTTTCTGATACCGTACACGAGACGACACTATC	161
QY	42	SeThrIleGlyValAspPheLysIleargThrIleGluLeuAspGlyLysTrhIleLys	61
DB	162	AGCACACATCGGGGTGAGACTTCAAGATCGAAGCATGACGAGTGGATGGCAAACTATCAA	221
QY	62	LeuglnIle-----	64
DB	222	CTTTCAGATCTGGGACACAGCGGGCCAGAACGGTTCGGACCATCTTCAGACTATAC	281
QY	65	-----GluSerPheAsn 69	
DB	282	CGGGGGGCTGATCGCATATCGTGCTATACCTGACCTGACACGAGATCTTACCGCAC	341
QY	70	ValLysTrpLeuGlnGluIleAspArgTyrAlaSerGlyAsnValAsnLysLeuLeu	89
DB	342	GAGAACGATGGGTGCGAGAGATGACCGCTATCCAGCAGACAGATGATACCTCTG	401
QY	90	ValGlyAsnLysCysAspLeuTrhTrhLysValValAspTyrTrhTrhAlaLysGlu	109
DB	402	GTGGGCAACGAGCGACTCTACACCAAGAGAGGTGGACACACACGACGACGAGG	461
QY	110	PheAlaAspSerLeuGlyIlePhePheLeuGluTrhSerAlaLysAsnAlaTrhAsnAl	129
DB	462	TTTGACATCTCTGGGATCTCCCTTTGGAGAGGAGCGCCAGAAATGCCACAAATGC	521
QY	130	GluGlnSerPheMetTrhMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThr	149
DB	522	GAGCAGCGCTGATGACCATGCGCTGTGAATCAAAAACGGGAGTGGGGCTGGACACCC	581
QY	150	AlaGlyIleValGluLysSerAsnValLysIleGlySerTrhProValLysGlnSerGly	169
DB	582	TTCTGGGGGCG---GAGCGGGCCCATCTCAAGATCGACAGCACGCCCTGTAAAGCGCGCTGC	638
QY	170	GlyGlyCysCys 173	
DB	639	GGTGGCTGTTC 650	
RESULT 13			
LOCUS	HSW801608	1985 bp	mRNA
DEFINITION	homo sapiens mRNA: cDNA DKFP5641172 (from clone DKFP5641172);		
ACCESSION	complete cds.		
VERSION	ALI36635		
KEYWORDS	ALI36635.1 GI:12052795		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Kumazawa, M. et al.: Chordata: Craniata; Vertebrata: Euteleostomi;		
AUTHORS	Edwards, J., Eulenski, Primates: Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1985)		
JOURNAL	Duesterhoft, A., Lauder, J., Mewes, H.W., Gassenhuber, J. and		
COMMENT	Wiemann, S.		
	Direct Submission		
	Submitted (12-MAR-2002) MIPS, Am Klopferspitze 18a, D-82152		
	Martinsried, GERMANY		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer		
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;		

sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp5641172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clonezrpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

1. 1985
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="11q13"
/clone="DKFZp5641172"
/tissue_type="brain"
/clone_lib="564 (synonym: hfbz2). Vector pAMP1; host X1-2blue; sites NotI + SalI"
/dev_stage="fetal"
1. 1985
/gene="DKFZp5641172"
48. 653
/gene="DKFZp5641172"
/note="strong similarity to Rab1"
/product="hypothetical protein"
/protein_id="CA86570.1"
/db_xref="GI:12052796"
/db_xref="SPTREMBL:Q9H0U4"
/translation="MNPEYDYLFLKLLIGSGVGSCLLRFADDTYTESYITGVD FKRIETLEDSKTIKLIOMDPAOGERRTTSSYVRGAHGIIYVDYTDOSYANVQW LQETIDRASENVKLVGNKSDITTKKYVDNTAKKEFADSLGTFLETSKKNATNVEQ AFETMAEIKRRKPGASGGERPNLKIDSTPVKASGGCC"
1860. 1865
/gene="DKFZp5641172"
1901
polya_signal
/gene="DKFZp5641172"
BASE COUNT 432 a 607 c 549 g 397 t
ORIGIN

Alignment Scores:

Pred. No.: 1.46e-65 Length: 1985
Score: 768.50 Matches: 156
Percent Similarity: 80.39% Conservative: 8
Best Local Similarity: 76.47% Mismatches: 7
Query Match: 86.74% Indels: 33
DB: Gaps: 2

US-09-820-003a-2 (1-173) x HSM801608 (1-1985)

QY 2 SerSerMetAsnProGluTYRAspTYRLeuPheLysLeuLeuIleGlyAspSerGly 21
Db 42 GCCGCCATGAACCCCGAATATGACTGTTTAAAGCTGCTTTGATTGGCGACTCAAGC 101
QY 22 ValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTYRThrGluSerTyrIle 41
Db 102 GTGGGCAAGTCATCGCTGCTCTGGGTTTGTCTGATGACAGTACAGAGAGCTACATC 161
QY 42 SerThrIleGlyValAlaAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61
Db 162 AGCAGCAATCGGGGTGACTTCAAGATCGAACAATCGAGCTGAGTGGCAAAACATATCAA 221
QY 62 LeuGlnIle----- 64
Db 222 CTTTCAGATCTGGACACAGCGGGCCAGGAACGGTTCGGACATCACTTCAGCTACTAC 281
QY 65 -----GluSerPheAsnAsn 69
Db 282 CGGGGGGCTCATGGCATCATCGTGGTGATGACCTCACTGACCAAGGAATCTTACGCCAAC 341
QY 70 ValIleGlnTrpLeuGlnIleValAspArgTYRAlaSerGluAsnValAlaLysLeuLeu 89
Db 342 GTGAAGCAGAGTGGCTGACGAGATTCAGCCGATCCAGCGAAGACGTCAATAACTCTCTG 401
QY 90 ValGlyAsnLysCysAspLeuThrThrLysLysValAlaAspTYRThrThrAlaLysGlu 109

Db 402 GTGGCAACAAGACCGACCTCACCAAGAGGTGTGGACCAACACCAAGCAGAG 461
QY 110 PheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAspVal 129
Db 462 TTTCAGACTCTCTGGGATCCCTTCTTGAGACGACGCCAAGAAATGCCAATATGTC 521
QY 130 GluInSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThr 149
Db 522 GAGCAGCGCTTCATGACCATGCTGTGTAATCAAAAGCGGATGGGCGCTGGAGCAGCC 581
QY 150 AlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
Db 582 TCTGGGGGCGC---GAGCGCGCCAAATTCGAAGATGCACACACCCCTGTAAAGCCGGCTGGC 638
QY 170 GlyLysCysCys 173
Db 639 GGTGGCTGTTC 650

RESULT 14

RNRAB1B
LOCUS RAB1B 654 bp mRNA linear ROD 12-SEP-1993
DEFINITION Rat cDNA for ras-related rab1b protein.
ACCESSION X13905
VERSION X13905.1 GI:57005
KEYWORDS GTP-binding protein; rab1b gene; rab1b protein; ras-related protein.

SOURCE

ORGANISM

Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

1 (Bases 1 to 654)
AUTHORS Touchot, N.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1989) Touchot N., Inserm U 248, Faculte de Medecine Lariboisiere St Louis, 10 Avenue de Verdun, 75010 Paris, France

REFERENCE

2 (Bases 1 to 654)
AUTHORS Vieh, E., Touchot, N., Zahraoui, A. and Tavillan, A.
TITLE Nucleotide sequence of a rat cDNA: rab1b, encoding a rab1-YPT related protein
JOURNAL Nucleic Acids Res. 17 (4), 1770 (1989)
MEDLINE 89160341
PUBMED 2493636

COMMENT

FEATURES

source

1. 654
/organism="Rattus sp."
/db_xref="taxon:10118"
/tissue_type="brain"
12. 617
/note="rab1b protein (AA 1 - 201)"
/codon_start=1
/protein_id="CAA32105.1"
/db_xref="GI:57006"
/db_xref="SWISS-PROT:P10536"
/translation="MNPEYDYLFLKLLIGSGVGSCLLRFADDTYTESYITGVD FKRIETLEDSKTIKLIOMDPAOGERRTTSSYVRGAHGIIYVDYTDOSYANVQW LQETIDRASENVKLVGNKSDITTKKYVDNTAKKEFADSLGTFLETSKKNATNVEQ AFETMAEIKRRKPGASGGERPNLKIDSTPVKASGGCC"

BASE COUNT 174 a 156 c 184 g 140 t
ORIGIN

Alignment Scores:

Pred. No.: 1.51e-65 Length: 654
Score: 762.50 Matches: 154
Percent Similarity: 79.90% Conservative: 9
Best Local Similarity: 75.49% Mismatches: 8
Query Match: 86.06% Indels: 33
DB: Gaps: 2

US-09-820-003a-2 (1-173) x RNRAB1B (1-654)

Mon Nov 18 08:42:03 2002

us-09-820-003a-2.rge

Page 13

[illegible]

Search completed: November 17, 2002, 13:58:04
Job time : 2672 secs

XX	Homo sapiens.
OS	
XX	
XX	W0200228999-A2.
XX	
XX	11-APR-2002.
PD	
XX	
XX	03-OCT-2001; 2001WO-US30821.
PF	
XX	03-OCT-2000; 2000US-237189P.
PR	
XX	(GENE-) GENE LOGIC INC.
PA	
XX	
XX	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
PI	
XX	WPI: 2002-435328/46.
DR	
XX	
PT	Detecting granulocyte activation by detecting differential expression
PT	of genes associated with granulocyte activation, which serves as
PT	diagnostic markers that is useful for monitoring disease states and
PT	drug toxicity
XX	
XX	Claim 1; SEQ ID No 549; 11app; English.
XX	
XX	The invention relates to detecting (M1) granulocyte (GC) activation
CC	(GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC	DNA chip analysis as given in the specification, and comparing
CC	the expression level to an expression level in an unactivated
CC	GC, where differential expression of Gs is indicative of GCA.
CC	Also included are modulating (M2) Gs by contacting GC with an agent
CC	that alters the expression of at least one gene in Gs; (2) screening (M3)
CC	for an agent capable of modulating GCA or an inflammation (especially
CC	chronic) in a tissue, an allergic response in a subject, exposure of a
CC	subject to a pathogen or sterile inflammatory disease using the
CC	gene expression profile; (3) detecting (M4) an inflammation (especially
CC	chronic) in a tissue, an allergic response in a subject, exposure of a
CC	subject to a pathogen or sterile inflammatory disease, by detecting the
CC	level of expression in a sample of the tissue of gene(s) from Gs, where
CC	the level of expression of the gene is indicative of inflammation;
CC	(4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC	an allergic response in a subject, exposure of a subject to a pathogen
CC	or sterile inflammatory disease, by contacting a tissue having
CC	inflammation with an agent that modulates the expression of gene(s)
CC	from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC	modulating Gs; M3 is useful for screening an agent capable of modulating
CC	GCA preferably in an inflammation in a tissue; M4 is useful for
CC	detecting an inflammation (especially chronic) in a tissue, an allergic
CC	response in a subject, exposure of a subject to a pathogen or sterile
CC	inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC	glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC	reperfusion injury, ARDS, adult respiratory distress syndrome,
CC	inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC	periodontal disease, also bacterial infection, viral infection,
CC	parasitic infection, protozoal infection, fungal infection and M5 is
CC	useful for treating one of the above conditions. The present
CC	sequence represents a gene differentially expressed in granulocytes.
CC	Note: The sequence data for this patent did not form part
CC	of the printed specification, but was obtained in electronic
CC	format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	Sequence 723 BP; 225 A; 142 C; 180 G; 176 T; 0 other;
XX	
XX	Alignment Scores:
Score:	4,33e-89 Length: 723
Percent Similarity:	860.00 Matches: 173
Best Local Similarity:	84.39% Conservative: 0
Query Match:	84.39% Mismatches: 0
DB:	97.07% Indels: 32
	24 Gaps: 1

Oy	1	MeSerSerMetAsnProGluTyrAspTyrLeuPheLeuLeuIleGlyAspSer	20
Db	51	ATGTCACAGCATGAAATCCCGAATATGATTTATTTATTCAGAGTTACTTTCGATTTGGGCACTCA	110
Oy	21	GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspSerThrTyrGluSerTyr	40
Db	111	GGGGTTGGAAAGTCCTGGCTCTCTTCTTAGGTTTCAGATGATCATATATACAGAAAGCTAC	170
Oy	41	IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle	60
Db	171	ATCAGCACAAATGGTGTTGATTTCAAAATAGAACTTATGAGTTAGACGGGAAAAACATC	230
Oy	61	LysLeuGluIle-----	64
Db	231	AAGCTTCAAAATATGGGACACAGCAGCCAGGAAGATTTCGAACATATCACCCTCCAGTTAT	290
Oy	65	-----GluSerPheAsn	68
Db	291	TACAGAGAGGCCCATGGCATCATGTTGTGTATGATGTGACAGATCAGAGAGTCCCTTCAT	350
Oy	69	AsnValLysGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu	88
Db	351	AATGTTAAACAGTGGCTGACAGAAATAGATGCTTATGCCACTGAAAAATGTCACAACAATTG	410
Oy	89	LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys	108
Db	411	TTGGTAGGGAACAATATGATCTACACCAAGAAAGACTAGTACTACACACAGCCAGC	470
Oy	109	GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn	128
Db	471	GAATTTGCTGATTCCTTGGAAATCCGTTTGTGAAACAGAGTGTAAAGATGCACAGCAAC	530
Oy	129	ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyVala	148
Db	531	GTAGAACAGTCTTTCATGACGATGCGAGCTGAGATTAAAAAGCAAAAGGGGTCCCGGAGCA	590
Oy	149	ThrAlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer	168
Db	591	ACAGTGGTGGTGTGCGAGAAAGTCCAATGTTAAAAATTCAGACGACTCCAGTCAACGACTCA	650
Oy	169	GlyGlyGlyCysCys 173	
Db	651	GGTGGAGGTTGCTGTC 665	
RESULT 2			
AC	AAC60009	AAC60009 standard; cDNA; 2528 BP.	
XX	AAC60009;		
XX	26-JAN-2001 (first entry)		
DE	Human secreted protein gene 44 SpQ ID NO:54.		
XX			
KW	Human; secreted protein; diagnosis; cyrostatic; immunosuppressive;		
KW	nocotropic; neuroprotective; antiviral; antiallergic; hepatotropic;		
KW	antidiabetic; antihistaminatory; antitumor; anticonvulsant;		
KW	antibacterial; antifungal; antiparasitic; cardiant; gene therapy;		
KW	cancer; immune disorder; cardiovascular disorder; wound healing;		
KW	neurological disease; infectious disease; chromosome identification; ss		
OS	Homo sapiens.		
XX			
XX	MO200058356-A1.		
XX	05-OCT-2000.		
XX			
PF	22-MAR-2000; 2000WO-US07535.		
XX			
PR	26-MAR-1999; 99US-0126511.		
PR	17-DEC-1999; 99US-0172413.		
XX			
PA	(HDMA-) HUMAN GENOME SCI INC.		

```

XX PI Rosen CA, Ruben SM, Komatsoulis G;
XX DR WPI: 2000-594639/56.
XX P-PSDB: AAB34816.
XX PT Fifty nucleic acid molecules encoding human secreted proteins, useful
XX PT in the prevention, treatment and diagnosis of cancer, immune disorders,
XX PT cardiovascular disorders and neurological diseases -
XX PS Claim 1: Page 371, 425pp; English.
XX CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the
XX CC human secreted proteins given in AAB34773 to AAB34822. AAB34823 to
XX CC AAB34832 represent human secreted polypeptide sequences and proteins
XX CC homologous to them, which are given in the exemplification of the present
XX CC invention. Human secreted proteins have activities based on the tissues
XX CC and cells the genes are expressed in. Examples of activities include:
XX CC cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
XX CC antiallergic; hepatotropic; antibacterial; antifungal; antiparasitic; and
XX CC cardiant. The polynucleotides and polypeptides are useful for
XX CC preventing, treating or ameliorating a medical condition in e.g. humans,
XX CC mice, rabbits, goats, dogs, cats, chickens or sheep. The
XX CC polypeptides can also be used as a food additive or preservative to
XX CC increase or decrease storage capabilities. The polynucleotides are
XX CC useful for chromosome identification. They are also useful as probes for
XX CC diagnosing a disorder related to the female reproductive system,
XX CC particularly breast and/or ovary cancer. They are also useful in the gene
XX CC therapy of breast and ovarian cancer. Nucleic acids, protein, antibodies,
XX CC agonists and antagonists from the present invention are useful in the
XX CC diagnosis, treatment and prevention of cancer, immune disorders,
XX CC cardiovascular disorders, wound healing, neurological diseases and
XX CC infectious disease. AAC59957 to AAC59965 and AAB34772 represents sequence
XX CC used in the exemplification of the present invention.
XX SQ Sequence 2528 BP; 772 A; 428 C; 545 G; 781 T; 2 other:

Alignment Scores:
Pred. No.: 2,24e-88 Length: 2528
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best Local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
DB: 21 Gaps: 1

US-09-820-003A-2 (1-173) x AAC60009 (1-2528)
OY 1 MetSerSerMetAspProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
DB 208 ATGTCACGACATGATCCGAAATGATTTATTCAGTACTCTGATTCGCGACTCA 267
OY 21 GlyValGlyLysSerCysLeuLeuArgPheAlaAspAspThrTyrThrcdLysSerTyr 40
DB 268 GGGGTGGAAAGCTTCCCTCTTACGCTTGCAGATGATACATATATACAAAGCTAC 327
OY 41 IleSerThrIleGlyValAspPheLysIleArgPheIleGluAspLysPheThrIle 60
DB 328 ATCCACCAATTCGTGCTGATTTCAAAATTAAGACTATAGGTTAAGCGGAAACAATC 367
OY 61 LysLeuGlnIle----- 64
DB 388 AAGCTTCAATTAATGAGACACAGCAGCGCAAGAAATTTGCAACAATCACTCCATTTAT 447
OY 65 -----GluserPheasn 68
DB 448 TACAGAGACCCATGCGATATAGTTGTGTATGATGTGACAGATCAGAGAGCTTCATAT 507
OY 69 AsnValLysGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
DB 508 AATGTTAAACAGGGCTGCAGGAAATGATCTTATGCGAGTAAATCAACAATATG 567
OY 89 LeuValGlyLysLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLys 108

```

```

DB 568 TTGATAGGACAAATATCTATCTGACCCAAAGAAAGTATGACTACAAACAGCAG 627
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
DB 628 GAAATTCGTGATTCCTTGGAATTCGGATTTTGGAAACGAGTGCCTAAAGAAATGCA 687
OY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
DB 688 GTAGAACAGCTCTTTCATGACGATGACGCTGAGATTAAACCAAGGAGTCCGAGACA 747
OY 149 ThrAlaGlyAlaAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
DB 748 ACAGCTGGTGGTGGTGAAGATCCAAATGTTAAATTCAGAGCTCCATCAGACAGTCA 807
OY 169 GlyGlyGlyCysCys 173
DB 808 GGTGGAGGTTCCTGC 822

RESULT 3
AAX27232
ID AAX27232 standard; DNA; 925 BP.
XX AAX27232;
XX 28-MAY-1999 (first entry)
XX Human Rab protein, RABP-2, coding sequence.
XX Rab protein; RABP-1; RABP-2; RABP-3; human; vesicle trafficking; cancer;
XX cell differentiation; apoptosis; immunodeficiency; cell proliferation;
XX neurodegenerative disease; myelodysplastic syndrome; wasting disease;
XX toxin-induced disease; infection; genetic defect; diagnosis; therapy; ss.
XX Homo sapiens.
XX MO9909182-A2.
XX 25-FEB-1999.
XX 17-AUG-1998; 98MO-US16983.
XX 21-AUG-1997; 97US-0916901.
XX (INCY-) INCYTE PHARM INC.
XX Corley NC, Hillman JL, Lal P, Shah P;
XX WPI: 1999-181042/15.
XX P-PSDB: AAY00919.
XX New purified human Rab proteins - used to develop products for
XX treating e.g. AIDS, immunodeficiencies, neurodegenerative diseases,
XX cancer, inflammation or autoimmune diseases.
XX Claim 29; Fig 2; 94pp; English.
XX This sequence encodes a human Rab protein of the invention, designated
XX RABP-2. The Rab proteins, RABP-1, RABP-2 and RABP-3 are involved in
XX vesicle trafficking, cell function, and cell differentiation. The RABP
XX polypeptides, DNAs and agonists can be used to prevent or treat a
XX disorder associated with an increase in apoptosis, e.g. infectious or
XX genetic immunodeficiencies, neurodegenerative diseases such as
XX Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis,
XX retinitis pigmentosa, and cerebellar degeneration, myelodysplastic
XX syndromes such as aplastic anemia, ischaemic injuries such as myocardial
XX infarction, stroke, and reperfusion injury, toxin-induced diseases such
XX as alcohol-induced liver damage, cirrhosis, and lathyrism, wasting
XX diseases such as cachexia, viral infections, and osteoporosis. They can
XX also be used to stimulate cell proliferation for use in transplantation
XX or to produce cells to fight an infection or a cancer or to correct a
XX genetic defect in a disease such as sickle cell beta thalassemia, cystic
XX fibrosis or Huntington's chorea. Antagonists can be used to prevent or

```

CC treat a disorder associated with cell proliferation e.g. cancers or
 CC inflammation, e.g. Addison's disease, adult respiratory distress
 CC syndrome, allergies, asthma, atherosclerosis, bronchitis, cholelithiasis,
 CC Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis,
 CC diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis,
 CC gout, Graves' disease, hyperostosis, irritable bowel syndrome, lupus
 CC erythematous, multiple sclerosis, myasthenia gravis, inflammation,
 CC osteoarthritis, osteoporosis, pancreatitis, polyomyelitis, rheumatoid
 CC arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis,
 CC complications of cancer, haemodialysis, extracorporeal circulation,
 CC infections and trauma.
 XX
 SQ Sequence 925 BP; 198 A; 263 C; 288 G; 176 T; 0 other;

Alignment Scores:
 Pred. No.: 1,78e-78 Length: 925
 Score: 768.50 Matches: 156
 Percent Similarity: 80.39% Conservative: 8
 Best Local Similarity: 76.47% Mismatches: 7
 Query Match: 86.74% Indels: 33
 DB: 20 Gaps: 2

US-09-820-003a-2 (1-173) x AAF2161 (1-925)

QY 2 SerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSergly 21
 Db 60 GCCCCCATGAAACCCGAAATATGACTGTTTAAGCTGTTTGAATGGCGACTAGCC 119
 QY 22 ValGlyLysSerGlyLeuLeuLeuArgPheAlaAspSeryTyrThrGluSerTyrIle 41
 Db 120 GTGGGCAAGTCATCGCTCTCTCGGTGCTGATGACAGTCACAGAGAGCTACATC 179
 QY 42 SerThrIleGlyValAlaPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61
 Db 180 AGCAGCATCGGGGGGACTTCATCATCGAAGCATCGAGCTGATGGCAAAACATATCAA 239
 QY 62 LeuGlnIle----- 64
 Db 240 CTTGATCTGGGACAGCGGGCCAGGAACGTTCCGACATCTCCAGCTACTAC 299
 QY 65 -----GluSerPheAsnAsn 69
 Db 300 CGGGGGCTCATGCGCATCATCTGTGTGATGACCTACTGACCGAAGATCTACGCCAAC 359
 QY 70 ValGlyGlnTyrLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeu 89
 Db 360 GTGAAGCAGTGGCTGACGAGGATGACCGCTATGCCAGCGAAGCTCAATTAAGCTCTCG 419
 QY 90 ValGlyAsnLysCysAspLeuThrThrLysLysValAlaAspTyrThrAlaLysGlu 109
 Db 420 GTGGGCAACAAAGAGCTTCACCAACAGAGTGTGGCAACACACCGACAGCCAGGAG 479
 QY 110 PheAlaAspSerLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsnVal 129
 Db 480 TTTCACACATCTCTGGGCAATCCCTTTGGAGACGAGCGCAAGATGCGACCAATGTC 539
 QY 130 GluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThr 149
 Db 540 GAGCAGCGCTTCATGACCATGAGCTGCGAATCAAAAAGGAGGGGCGCTGGAGCAGCC 599
 QY 150 AlaGlyLysAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSergly 169
 Db 600 TCTGGGGGCGC--GAGCGCCCAATCTCAAGATCGACAGCACCCTGTAAGCCGCGTGC 656
 QY 170 GlyGlyCysCys 173
 Db 657 GGtGGCTGTTC 668

RESULT 4
 AAF2161
 ID AAF2161 standard; DNA: 939 BP.
 XX AAF2161;
 AC AAF2161;
 4

XX 27-MAR-2001 (first entry)
 DT Human breast and ovarian cancer associated antigen gene SEQ ID 48.
 XX
 DE

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neotropic; neuroprotective; antiviral; antifungal; hepatotropic;
 KW antidiabetic; antifungal; antiviral; antiviral; antiviral; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease; ds.
 XX
 OS Homo sapiens.

XX NO200055173-A1.
 XX 21-SEP-2000.
 XX 08-MAR-2000; 2000MO-US05881.
 XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 XX WPI, 2000-611515/58.
 XX P-PSDB; AAB58758.

XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -
 XX
 PS Claim 1; Page 515; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neotropic; neuroprotective; antiviral; antifungal; hepatotropic;
 CC antidiabetic; antiinflammatory; antitumor; vulnery; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiac activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemia; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX

SQ Sequence 939 BP; 202 A; 264 C; 289 G; 180 T; 4 other;

Alignment Scores:
 Pred. No.: 1,82e-78 Length: 939
 Score: 768.50 Matches: 156
 Percent Similarity: 80.39% Conservative: 8
 Best Local Similarity: 76.47% Mismatches: 7
 Query Match: 86.74% Indels: 33
 DB: 21 Gaps: 2

US-09-820-003a-2 (1-173) x AAF2161 (1-939)

QY 2 SerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSergly 21
 Db 64 GCCGCCATGAAACCCGAAATATGACTGTTTAAGCTGTTTGAATGGCGACTAGCC 123

Db 166 AGCACATCGGGGTGGACTTCAGATCCGAAACCATCGAGCTGGATGGCAAACTATCAAA 225
Qy 62 LeuGlnIle----- 64
Db 226 CTTGAGATCTGGACACAGCGGGCCAGAGCGTTCCGACCATTCACTTCAGCTACTAC 285
Qy 65 -----GlutSerPheAsnAsn 69
Db 286 CGGGGGGCTCATGTCATCATGCTGTATGACGTCACCTGACCAAGAACTCTACGCCAAC 345
Qy 70 ValIysGlnTrpLeuGlnGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeu 89
Db 346 GTGAAGCAGTGGCTGCAGAGATTGACCGCTATGCGACGCGAAGCTCAATTAAGCTCTG 405
Qy 90 ValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLysGlu 109
Db 406 GTGGGCAACAGACGACCTCACCACCAAGAGGTGGTGAGACACACACAGCCAAAGAG 465
Qy 110 PheAlaAspSerLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsnVal 129
Db 466 TTTGACAGACTCTGCGGATCCCTCTTGGAGACGAGCGCCAGAAATGCCACCAATGTC 525
Qy 130 GluInserPheMetThrMetAlaAlaGlnIleLysLysArgMetGlyProGlyAlaThr 149
Db 526 GAGCAGGCGTTCATGACCATGCTGCTGAATCAAAAAGCGATGGGCTCGAGACACCC 585
Qy 150 AlaGlyAlaGlnLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
Db 586 TCTGGGGGCG---GAGCGGCCCAATCTCAAGATCGACACACCCCTGTAAAGCCGGCTGCG 642
Qy 170 GlyGlyCysCys 173
Db 643 GGTGGCTGTTCG 654
RESULT 6
AA193456
ID AA193456 standard; cDNA; 1944 BP.
XX
AC AA193456;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 13516.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PE 26-FEB-2001; 2001MO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
DR P-PSDB; AA013525.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukemia, inflammation and immune
PT disorders -
XX
PS Claim 1; SEQ ID NO 13516; 1399pp + Sequence Listing; English.
XX

CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
SQ Sequence 1944 BP; 372 A; 618 C; 552 G; 402 T; 0 other;
Alignment Scores:
Pred. NO.: 4.72e-78 Length: 1944
Score: 768.50 Matches: 156
Percent Similarity: 80.39% Conservative: 8
Best Local Similarity: 76.47% Mismatches: 7
Query Match: 86.74% Indels: 33
DB: 22 gaps: 2
US-09-820-003a-2 (1-173) x AA193456 (1-1944)
Qy 2 SerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGly 21
Db 56 GCGCGCATGAACCCGGAATATGACTACCTGTTAAGCGCTTTGATGGGAGTCAAGCC 115
Qy 22 ValGlyLysSerCysLeuLeuLeuArgPheAlaAspThrTyrThrGluSerTyrIle 41
Db 116 GTGGGCAAGTCATCCTGCTCTCTCGGTTTGTCTATGACAGTACAGAGAGCTACATC 175
Qy 42 SerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61
Db 176 AGCACCATCGGGGTGGACTTCAGATCGAAGCATCGAGCTGGATGGCAAAACTATCAAA 235
Qy 62 LeuGlnIle----- 64
Db 236 CTTGAGATCTGGACACAGCGGGCCAGAGACGTTCCGACCATCACTTCAGCTACTAC 295
Qy 65 -----GlutSerPheAsnAsn 69
Db 296 CGGGGGGCTCATGTCATCATGCTGTATGACGTCACCTGACCAAGAACTCTACGCCAAC 355
Qy 70 ValIysGlnTrpLeuGlnGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeu 89
Db 356 GTGAAGCAGTGGCTGCAGAGATTGACCGCTATGCCAGCAAGCTCAATTAAGCTCTG 415
Qy 90 ValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrAlaLysGlu 109
Db 416 GTGGGCAACAGACGACCTCACCACCAAGAGTGGTGAGCAACACACAGCCAAAGAG 475
Qy 110 PheAlaAspSerLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsnVal 129
Db 476 TTTGACAGACTCTGCGGATCCCTCTTGGAGACGAGCGCCAAAGAAATGCCACCAATGTC 535
Qy 130 GluInserPheMetThrMetAlaAlaGlnIleLysLysArgMetGlyProGlyAlaThr 149
Db 536 GAGCAGGCGTTCATGACCATGCTGCTGAATCAAAAAGCGATGGGCTCGAGACGCC 595
Qy 150 AlaGlyAlaGlnLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
Db 596 TCTGGGGGCG---GAGCGGCCCAATCTCAAGATCGACAGCACCCCTGTAAAGCCGGCTGCG 652
Qy 170 GlyGlyCysCys 173
Db 653 GGTGGCTGTTCG 664
RESULT 7
AAS34668

ID AAS34668 standard; DNA; 1965 BP.
XX AAS34668;
AC
XX
XX 17-DEC-2001 (first entry)
DT
XX
XX Human DNA for a novel foetal antigen, SEQ ID No 2092.
DE
XX
XX Human: foetal tissue antigen; ds; antiinflammatory; neuroprotective;
KW immunomodulator; haemolysate; cytosolic; nephrotoxic;
KW cardiovascular; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; breast neoplasm; cancer;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; food additive.
XX
XX Homo sapiens.
XX
XX WO20015312-A2.
PN
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01321.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189876.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-020515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239937.
PR 13-OCT-2000; 2000US-0239938.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241825.
PR 01-NOV-2000; 2000US-0244826.
PR 08-NOV-2000; 2000US-0244874.
PR 08-NOV-2000; 2000US-0244875.
PR 08-NOV-2000; 2000US-0244876.
PR 08-NOV-2000; 2000US-0244877.
PR 08-NOV-2000; 2000US-0244878.
PR 08-NOV-2000; 2000US-0245223.
PR 08-NOV-2000; 2000US-0245224.
PR 08-NOV-2000; 2000US-0245225.
PR 08-NOV-2000; 2000US-0245226.
PR 08-NOV-2000; 2000US-0245227.
PR 08-NOV-2000; 2000US-0245228.
PR 08-NOV-2000; 2000US-0245328.
PR 08-NOV-2000; 2000US-0245329.
PR 08-NOV-2000; 2000US-0245330.
PR 08-NOV-2000; 2000US-0245610.
PR 08-NOV-2000; 2000US-0245611.
PR 08-NOV-2000; 2000US-0245613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

Db 180 TCCATCATGACCCCTGCAATCGACTATCTCTTACAGCTCCTTCTTATCGAGATCCGGT 239
QY 22 ValGlyLysSerCysLeuLeuLeuArgPheAlaAspThrTyrThrGluSerTyrIle 41
Db 240 GTTGAAATAATCTTGCTGCTACTCTCGCTTGCGAGACAGACACCTACAGAGAGCTATATC 299
QY 42 SerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61
Db 300 TCCACTATGTGCTGTGATTTAAATCCGAACATCACTGATGTCAGACAGACAGTGAAG 359
QY 62 LeuGlnIle----- 64
Db 360 CTTTCAGATTGGGACACTGCGGGCCAGAGCGGTTCCGACACCATCAGCTGCTTACTAT 419
QY 65 -----gluSerPheAsnAsn 69
Db 420 CGAGGTCTCATGTATCTGTGCTGATGATGTTACTGATGATGATGATGATGATGATGAT 479
QY 70 ValLysGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeu 89
Db 480 GTGAAAGAGTGGCTCCAGGAGATCGATCGCTATGCCACTAGCGGTGTCAACAGCTGCTT 539
QY 90 ValGlyAsnLysCysAspIleuThrThrLysLysValValAspTyrThrThrAlaLysGlu 109
Db 540 GTGGGTACACAGAGTGCACATGAAGATAAAGGTCTGGAGTACACGCTGGCCAAAGGAG 599
QY 110 PheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnVal 129
Db 600 TTGGCTGATGACCTTGGAATATACCTTCTTGAGACCTCTGCTAAGATGCTCGAAGCTC 659
QY 130 GluGlnSerPheMetThrMetAlaIleGluIleLysLysArgMetGlyProGlyAlaThr 149
Db 660 GAGCAAGACCTTCTTGACAAAGGCAAGGACATCAAGGAGGTATGGGT-----ACC 710
QY 150 AlAGlyValAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
Db 711 GCCACTGTCAACACAGCCGACTGTGACAGTGGCCAGGCCAGGGTGTCCAGTCTGGG 770
QY 170 -----GlyGlyCysCys 173
Db 771 TCCGACAGGTGTGCTGC 788
AC 33987;
AC 33987
ID AAC33987 standard; DNA; 959 BP.
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5039.
XX
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

Db 714 GCAGGAGGATCTAAGCCACCAACGCTCCAGATCCGAGACACCCGTTAACGACAA--- 770
OY 170 GYGYCysCys 173
Db 771 TCAGGCTGCTGC 782
RESULT 11
AAC35200
ID AAC35200 standard; DNA; 932 BP.
AC AAC35200;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9347.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132486.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144614.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148564.
 PR 16-AUG-1999; 99US-0149358.
 PR 17-AUG-1999; 99US-0149115.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157773.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159844.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 5,83e-63
 Score: 633.00
 Percent Similarity: 70.79%
 Best Local Similarity: 65.84%
 Query Match: 71.44%
 DB: 21
 Gaps: 3

US-09-820-003a-2 (1-173) x AAC35200 (1-932)

OY 4 MetAspProGluTyrAspTyrLeuPheLeuLeuLeuIleGlyAspSerGlyValGly 23
 DB 125 ATGAATCCTGAAATATGACTATCTGTCACCTCTGCTATTGGTGTGATCTGTTGCA 184
 OY 24 LysSerCysLeuLeuLeuArgPheAlaAspPheTyrThrGluSerThrIleSerThr 43
 DB 185 AAATATCCTGCTCTTCTAAGATTGCTGATGATCTTACCGGATACATCAATCAAGACC 244
 OY 44 IleGlyValAspPheLeuIleArgThrIleGluLeuAspGlyLysThrIleLeuGln 63
 DB 245 ATTGCTGTGTACTTTAAATTCGACAGTTCGACGAGGAGGAGGAGGAGGAGGAGGAG 304
 OY 64 Ile----- 64
 DB 305 ATCTGGACACAGCAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 364
 OY 65 -----GluSerPheAsnAsnValLys 71
 DB 365 GCTCAWGGATCATCTGCTACTTATGATGTCACAGACTGACAGACTTCAACACCTCAAA 424
 OY 72 GlnTPLeuGlnGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeuValGly 91
 DB 425 CATGGCTGATGAAATGACCGTACCGTACGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 484
 OY 92 AsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLysGluPheAla 111
 DB 485 AACAGAACGATCTCATCTTCACAGAAAGTTGATCCACGACGACGAGGAGGAGGAGGAGGAG 544
 OY 112 AspSerLeuGlyIlePhePheLeuGluThrSerAlaLysAsnAlaThrAsnValGlnGln 131
 DB 545 GATGAACTGTGGATCCCATTTCTGAAACAGAGTCTAAATAATGACAAACGATGAGAA 604
 OY 132 SerPheMetPheAlaAlaGluIleLysLysArgMetGlyProGlyAlaThrLysGly 151
 DB 605 GCTTCATGCTGATGACTCTGCTGCAATTAACAAAGATGCT---AGCCACACACTGCA 661
 OY 152 GlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGlyGly 171
 DB 662 GGTGCACAGCCACCAACGCTCAGATCGTGGACAGCAGTGAACGACAA---TCAGGC 718
 OY 172 CysCys 173
 DB 719 TGTGT 724
 DB 719 TGTGT 724
 RESULT 12
 AAS83860/c
 ID AAS83860 standard; cDNA, 1193 BP.
 AC AAS83860;
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #19664.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;

DR WPI, 2001-639362/73.
DR P-PSDB; ABG19673.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 19664; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1193 BP; 372 A; 279 C; 213 G; 329 T; 0 other;

Alignment Scores:
Pred. No.: 8.64e-62 Length: 1193
Score: 624.00 Matches: 129
Percent Similarity: 74.57% Conservative: 0
Best Local Similarity: 74.57% Mismatches: 0
Query Match: 70.43% Indels: 44
DB: 23 Gaps: 1

US-09-820-003a-2 (1-173) x AAS83860 (1-1193)
QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20
DB 1003 ATGCCGACGATGCAATCCGAAATGATTATTATTCACAACTTCTCGATTGGCGACTCA 944
QY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40
DB 943 GGGGTTGGAAAGCTTCCTCTCTTACGTTGACAGATACATATACAGAAAGCTAC 884
QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 883 ATCAGCACAATGGTGTGGATTTCACAAATAGACTATAGAGTAGACGGGAAACATC 824
QY 61 LysLeuGlnIleGluSerPheAsnAsnValLysGlnIlePheGlnGlnIleAspArgTyr 80
DB 823 AAGCTTCAATA----- 812
QY 81 AlaSerGluAsnValAsnLysLeuLeuValGlyAsnLysCysAspLeuThrThrLysLys 100
DB 812 ----- 812
QY 101 ValValAspTyrThrThrAlaLysGluPheAlaAspSerLeuGlyIleProPheLeuGlu 120
DB 811 -----GAAATTGCGGATTCCTTGGCAATTCGCTTTTGGAA 776
QY 121 ThrSerAlaLysAsnAlaThrAsnValGluGlnSerPheMetThrMetAlaAlaGluIle 140
DB 775 ACCAGTGCTTAAGATGCAAGCAATGTAAGACAGCTTTCATGACGATGGACGCTGACATT 716
QY 141 LysLysArgMetGlyProGlyAlaThrAlaGlyGlyValGluLysSerAsnValLysIle 160

DB 715 AAAAGCGAATGGGTCCCGGAGACACAGCTGTGTGCTGAGAACTTAATTT 656
QY 161 GlnSerThrProValLysGlnSerGlyGlyCysCys 173
DB 655 CAGAGCACTCCAGTCAGCAGCTAGGTGAGGTTGCTGC 617

RESULT 13
AAC42684
ID AAC42684 standard; DNA: 777 BP.
XX
AC AAC42684:
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36471.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 28-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 23-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.


```
DB: 24 Gaps: 2
US-09-820-003a-2 (1-173) x ABN98761 (1-881)

OY 3 SerMetAsnProGluTyrAspTyrLeuPheValLeuLeuIleGlyAspSerGlyVal 22
   |||||
DB 57 ACCATGATCTCTGATGACGATATCTTTTCACGCTCTGCTTATCGGGGATTCGGCTTA 116
OY 23 GlyLysSerCysLeuLeuPheAlaAspPheThrTyrThrGluSerTyrIleSer 42
   |||||
DB 117 GGCAGCTTCTCTCTCTTTCGATTCCTGATTCCTTATGATGACAACTTACATTCAT 176
OY 43 ThrIleGlyValAspPheValLeuGlyThrIleGluLeuAspGlyLysThrIleLeu 62
   |||||
DB 177 ACTATTGACGATGATTTTAAATTAAGNCTGTGAAACAGATGCAACATTAATAGCTC 236
OY 63 GlnIle----- 64
DB 237 CAAATTTGGACACTGCTGCTGACAGACGTTGACGACTTACTAGCAGTTACTACCGT 296
OY 65 -----GluSerPheAsnVal 70
DB 297 GGGGCACATGCAATTTATTTGCTACAGATGTCACAGATGAAGAACTCATATATGTC 356
OY 71 LysGlnTTPLeuGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeuVal 90
DB 357 AAGCATGCTTGATGATAATGATGCTTATGATGACATGCAACAACTCTTCTTT 416
OY 91 GlyAsnLysCysAspLeuThrThrLysValValAspTyrThrThrAlaLysGluPhe 110
   |||||
DB 417 GGAACAAAGTGTATCTTACTGAAACAGCCATCTTGTGAACCTGCAAGGCTTTT 476
OY 111 AlaAspSerLeuGlyIleProPheLeuGlnThrSerAlaLysAsnAlaThrAsnValGlu 130
   |||||
DB 477 GCCGATGAATACGGGATTCCTTTATGGAGACATGACAAAGATGCTCAAACTATAGA 536
OY 131 GlnSerPheMetThrMetAlaIleLysLysArgMetGlyProGlyAlaThrAla 150
   |||||
DB 537 CAGCCTTCATGCGAATGCTGTCATCATCAAGAGAGATGCTGACCAACACCTGGG 596
OY 151 GlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 170
   |||||
DB 597 AATAATGCAAGACCCAGCAGCCTGAGATCAGAGACAGCAGCCTGTGGCAGCAAG--AAC 653
OY 171 GlyCysCys 173
   |||||
DB 654 GGCTGCTGC 662

RESULT 15
AAC42764
ID AAC42764 standard; DNA; 666 BP.
XX
AC AAC42764;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36762.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
XX
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 05-MAR-1999; 99US-0123348.
PR 15-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 26-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 04-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134421.
PR 14-MAY-1999; 99US-0134570.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140603.
PR 24-JUN-1999; 99US-0140823.
PR 26-JUN-1999; 99US-0140861.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141642.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142855.
PR 06-JUL-1999; 99US-0142850.
PR 08-JUL-1999; 99US-0142820.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142817.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
```

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147302.
PR 04-AUG-1999; 99US-0147304.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159296.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Prod. No.:	9-5e-61	Length:	666
Score:	612.00	Matches:	128
Percent Similarity:	70.96%	Conservative:	11
Best Local Similarity:	64.97%	Mismatches:	24
Query Match:	69.07%	Indels:	34
DB:	21	Gaps:	3

US-09-820-003a-2 (1-173) x AAC42764 (1-666)

QY 9 Aspyrleuphelsleuleuleuileglyaspserglyvalglylyserscysleuleu 28
|||
DB 73 GACTATTGTTCAAGCTTCTGCTCATGCGTATTCGTGGAAAGTGGTCTGCT 132
|||
QY 29 Leuargpheaiaaspserthtyrthgluserthyrlleserthrlleiglyvalasphe 48
|||
DB 133 CTAAGTTGCTGATGATTCCTACCTGATGCTACATCACCATTCAGATTGGGATTCG 192
|||
QY 49 Lysileargthrillegluleaspglylvysrthrlleysleuglnile----- 64
|||
DB 193 AAAATCCGACAGTCGACAGCAAGATGAGAACCATCAATCCAGATTGGGATTCGCA 252
|||
QY 64 ----- 64
|||
DB 253 GCCCAGAGACGTTCCAGAACGATTACTAGCAGTTACTACAGAGAGCTCATGAATTATT 312
|||
QY 65 -----Gluserpheasnsnvallysglnthpplleuglnlu 76
|||
DB 313 GTGACTTATGATGTACAGATCTAGAAAAGCTTCAACACGTCAGCATGGCTTAATGAA 372
|||
QY 77 IleaspargytralisersgluasnvalasnllysleuleuvalglasnllysCysAspleu 96
|||
DB 373 ATCGACCGCTATGCAAGTGAGAGATGTTBACAGCTACTGCTGGAGAACAGTGTGATCTC 432
|||
QY 97 Thrthrlslyslvalaslyaspserthtyrthrlleaslysluphealaaspsersleuglylle 116
|||
DB 433 ACATCACAGAAAGTGTATCCACGACGACACTAAGGCTTGGCTGTATGAACCTGGGATC 492
|||
QY 117 PropheluengluthrseralalysasnalaThrasnValgluglnserPheMetThmet 136
|||
DB 493 CCATTCTTGAAACAGTGTAAAGATGCTACCATGTCCGAAAGAGCTTCATGCGCATG 552
|||

MON NOV 18 08:42:04 2002

us-09-820-003a-2.rng

Page 19

Search completed: November 17, 2002, 13:13:16
Job time : 259 secs

US-09-820-003a-2 (1-173) x US-09-820-003a-1 (1-1405)

```

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20
    |||
Db 180 ATGTCCGACATGAAATCCGAAATGATATTATTATTCAGTACTTCTTGATGGGACTCA 239
QY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40
    |||
Db 240 GGGGTGGAAAGTCTTCTCTTCTAGAGTTGACAGATGATACATATACAGAAAGCTAC 299
QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
    |||
Db 300 ATGAGCACAATGTTGGTGGATTTCATAAATAAGAACTATAGACGGGAAAAACATC 359
QY 61 LysLeuGlnIleGluSerPheAsnAsnValLysGlnIlePheGlnGlnIleAspArgTyr 80
    |||
Db 360 AACGTTCAATAGAGTCTCTTCATTAATGTTAAACAGTGGCTGCCAGAAATAGATCGTTAT 419
QY 81 AlaSerGluAsnValAsnLysLeuLeuValGlyAsnLysCysAspLeuThrTyrLysLys 100
    |||
Db 420 GCCAGTGAATAATGTCACAAATTTGTTGGTGGAGAACAAATGTGATCTGACACAAAGAAA 479
QY 101 ValValAspTyrThrThrAlaLysGluPheAlaAspSerLeuGlyIleProPheLeuGlu 120
    |||
Db 480 GTAGTACACTACACACAGGAGGAATTTGCTGATTCCTTGGAATTCGCTTTTGGAA 539
QY 121 ThrSerAlaLysAsnAlaThrAsnValGlnGlnSerPheMetThrMetAlaIleGlu 140
    |||
Db 540 ACCAGTCTAAGATGCAAGCAAGCAAGTGTACAGCTTTCATGACAGACGACGCAATAT 599
QY 141 LysLysArgMetGlyProGlyAlaThrAlaGlyLysIleGluLysSerAsnValLysIle 160
    |||
Db 600 AAAAAGGAATGGTCCCGGAGCAACAGCTGTGCTCTGAGAAGTCCAATGTTAAAT 659
QY 161 GlnSerThrProValLysGlnSerGlyGlyCysCys 173
    |||
Db 660 CAGAGCCTCAGTCAAGCAGTCAAGTGGAGGTTGCTGC 698

```

RESULT 2

US-09-967-736-4
 ; Sequence 4, Application US/09967736
 ; Patent No. US20020103340A1

GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.

LaI, Preeti
 Corley, Neil C.

Shah, Purvi

TITLE OF INVENTION: RAB PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTED for Windows Version 2.0

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/967,736

FILING DATE: 28-Sep-2001

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/154,602

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0367 US

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 925 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
 LIBRARY: LIVERU04

CLONE: 2514506

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-967-736-4

Alignment Scores:
 Pred. No.: 1,25e-91 Length: 925
 Score: 768.50 Matches: 156
 Percent Similarity: 80.39% Conservative: 8
 Best Local Similarity: 76.47% Mismatches: 7
 Query Match: 86.74% Indels: 33
 DB: 10 Gaps: 2

US-09-820-003a-2 (1-173) x US-09-967-736-4 (1-925)

```

QY 2 SerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSerGly 21
    |||
Db 60 GCCGCCATGAACCCCGAATATGACTACCTGTTAAAGCTGCTTTGATTGGCGACTCAGGC 119
QY 22 ValGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyrIle 41
    |||
Db 120 GTGGGCAAGTCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 179
QY 42 SerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLys 61
    |||
Db 180 AGCACCATCGGGGTGAGCTTCAAGATCGAACCATCGAGCTGATGCAAAACATATCAA 229
QY 62 LeuGlnIle----- 64
    |||
Db 240 CTTGAGATCTGGGACACAGCGGGCCAGAAAGGTTCCGACCATCACTTCCAGCTACTAC 299
QY 65 -----GluSerPheAsnAsn 69
    |||
Db 300 CGGGGGCTCATGATGATCATCGTGTATGACGTCACTGACCAAGGAATCTTACGCCAAC 359
QY 70 ValLysGlnIlePheGlnGlnIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeu 89
    |||
Db 360 GTGAAGCAGTGGCTGCGAGGATGACCGCTATGCCAGCAAGCAATTAAGCTCTCTG 419
QY 90 ValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLysGlu 109
    |||
Db 420 GTGGGCAACAAAGACGCTCACCCACCAACAGAGGTGTGACAAACACACACCAAGAG 479
QY 110 PheAlaAspSerLeuGlnIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnVal 129
    |||
Db 480 TTTCACACTCTCTGGGCAATCCCTTTTGGAGACGAGGCCCAAGAAATGCCACCAATGTC 539
QY 130 GluGlnSerPheMetThrMetAlaIleGluIleLysLysArgMetGlyProGlyAlaThr 149
    |||
Db 540 GAGCAGCGCTTCAATGACATGAGCGGTGTAATCAAAAAGGAGATGGGGCTTGAGCGCC 599
QY 150 AlaGlyLysAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGly 169
    |||
Db 600 TCTGGGGGCG---GAGCGGCCCAATCTCAAGATGACAGCACCCCTGTAAAGCCGGCTGC 656
QY 170 GlyGlyCysCys 173
    |||
Db 657 GTGGCTGTGTC 668

```

RESULT 3

US-09-938-842A-832
 ; Sequence 832, Application US/09938842A

; Patent No. US20020160378A1

GENERAL INFORMATION:
 APPLICANT: Harper, Jeff


```

QY 108 ----- 108
Db 182 TATGGGTTTCGACGTACAGTAAGCCACAGCCTTTAAAAATATGTGCATACAGTAATCT 241
QY 108 ----- 108
Db 242 GTGACATGACAAATTTGTGTAGCATCTGTTTGATCCATGAAGAACTTATGCTCAGCCT 301
QY 108 ----- 108
Db 302 CATTATGATGTAGAAATTCAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 361
QY 109 ----- 122
Db 362 GCTTCATTAATCTCTTACGAAATTTGCGATCTCCCTTGAATCCGTTTGGAAACCAAGT 421
QY 123 AAlaYsAsnAlaThrAsnValGluGlnSerPheMetThrMetAlaAlaGluLeuLys 142
Db 422 GCTAAGATGCAACGAATGTAGAAAGAGTCTTTCATGACGATGGACCTGAGATTAAAAAG 481
QY 143 ArgMetGlyProGlyAlaThrAlaGlyAlaGluLysSerAsnValLysIleGlnSer 162
Db 482 CGAATGGGTCCCGGAGCAACAGCTGGTGGTCTGAGAAAGTCCAAATGTTAAATTCAGAGC 541
QY 163 ThrProValLysGlnSerGlyGlyCys 173
Db 542 ACTCAGCTCAAGACTCAGCTGAGGTGCTGCTGC 574

RESULT 7
US-09-938-842A-836
; Sequence 836, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 836
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-836

Alignment Scores:
Pred. No.: 2,76e-47 Length: 651
Score: 428.50 Matches: 92
Percent Similarity: 58.85% Conservative: 31
Best Local Similarity: 44.02% Mismatches: 39
Query Match: 48.36% Indels: 47
DB: 9 Gaps: 5

US-09-820-003a-2 (1-173) x US-09-938-842A-836 (1-651)
QY 7 GluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGlyLysSerCys 26
Db 31 GATTACCATTAACCTTAACCTTCTGCTGATCGACACAGCCGCTTGTGAAGATTGC 90
QY 27 LeuLeuLeuArgPheAlaAspAspThrTyrThrGlnSerTyrIleSerThrIleGlyVal 46
Db 91 CTTCTTACGTTCTCAGATGGCTGTTTACACACAGTTTCATTAACAATATGGGATT 150

```

```

QY 47 AspPheLysIleArgThrIleGluLeuAspGlyThrIleLysLeuGlnLe----- 64
Db 151 GATTTTAAGATGAGACTATTAAGTGTAGGAGAGAGATTAAAGCTGCGAATCTGGAT 210
QY 64 ----- 64
Db 211 ACTGCCGACAGAGCGGTTCCGCAATACATCACTGCTACTACCGTGAGCCATGGG 270
QY 65 ----- 74
Db 271 ATTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330
QY 75 GlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeuValGlyAsnLysCys 94
Db 331 CTTAATGATGAGAGATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 390
QY 95 AspLeu---ThrThrLysLysValAlaAspTyrThrThrAlaLysGluPheAlaAspSer 113
Db 391 GATATGATGAAAGCAAAAGAGCTGCTCCAAATCTAAGGCGCAAGCTCTTCAATGAA 450
QY 114 LeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnValGluGlnSerPhe 133
Db 451 TATGGAATGAAGATTTCGAGACTAGTCCAAAGACTTAACCTTAACGTTGAGAGATTTC 510
QY 134 MetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAlaThrAlaGlyValAla 153
Db 511 TTCTCTATGCTAAAGACTTATACAAAGACTTCGAGAT-----ACCGATCAGAGCT 564
QY 154 GluLysSerAsnValLysIleGlnSerThrProValLysGlnSerGlyGly----- 171
Db 565 GAGCCGCAACATCAAAATTCACCAATCC-----GACCAGGTGGCGGAACATCT 615
QY 172 ----- CysCys 173
Db 616 CAGGCTACTCAGAAATCAGCAATGTTGC 642

RESULT 8
US-09-794-257-9
; Sequence 9, Application US/09794257
; Patent No. US2002009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Human, G-Proteins
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US2002009804A1
; FILE REFERENCE: 33800/209285
; CURRENT APPLICATION NUMBER: US/09/794, 257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185, 606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 624
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-794-257-9

Alignment Scores:
Pred. No.: 2,06e-45 Length: 624
Score: 414.00 Matches: 87
Percent Similarity: 60.50% Conservative: 34
Best Local Similarity: 43.50% Mismatches: 41
Query Match: 46.73% Indels: 39
DB: 10 Gaps: 3

US-09-820-003a-2 (1-173) x US-09-794-257-9 (1-624)
QY 4 MetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGly 23
Db 1 ATGGGGAAGAGATGATTAATCTTCAACCTCTGCTGATCGGCGAGCTGGGGGTAGGC 60
QY 24 LysSerGlyLeuLeuLeuArgPheAlaAspAspThrTyrThrGlnSerTyrIleSerThr 43

```

```
D 61 AAGACCTGCTCCGTGTCCTCCCTTCTCAGAGAGCGCTTCAACACCACCTTCATCTCCACC 120
Q 44 ILeGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLysLeuGln 63
D 121 ATGGGATTCGATTTTAAATTAAGACATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 180
Q 64 ILe----- 64
D 181 ATATGGGACACAGCGGGTCCAGAAAGATTCGAAACATCAAGACAGCGTACTACAGAGA 240
Q 65 -----GluSerPheAsnAsnValLys 71
D 241 GCCATGGGCAATTATGCTGCTGTATGACATCAAAATGAAAAATCTTGACAAATATTAA 300
Q 72 GlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuValGly 91
D 301 AATTGGATCAGAAACATTGAAAGCATGCCCTCTCCGATGTCGAAAGAAATGATCTGGGT 360
Q 92 AsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLysGluPheAla 111
D 361 AACAAATGTATGATGATGACAAAGAACAAAGTGTCAAAAGAAAGAGGAGGAGAACTAGCA 420
Q 112 AspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnValGluGln 131
D 421 ATTGACTATGGGATTAAATTCTTGAGACAAGCGCAAAATTCAGTGCAAATGTAGAAAGAG 480
Q 132 SerPheMetThrMetAlaIleGluIle-----LysLysArgMetGlyProGly 147
D 481 GCATTTTTCACCTTGACGACGAGATTAATGACAAACTCAACAGAAAAATGAATGACAGC 540
Q 148 AlaThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGln 167
D 541 AATTACACGAGAGAGAGCT-----GACACAGTGAATAATACAGAAA-CCGATCAAGAA 593
```

RESULT 9

```
US-09-794-257-7
; Sequence 7, Application US/09794257
; Patent No. US2002009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US2002009804A1el
; TITLE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (18)...(641)
US-09-794-257-7
```

Alignment Scores:

```
Score: 5.03e-45 Length: 1161
Percent Similarity: 414.00 Matches: 87
Best Local Similarity: 60.508 Conservative: 34
Query Match: 43.508 Mismatches: 41
46.738 Indels: 39
10 Gaps: 3
```

US-09-820-003a-2 (1-173) x US-09-794-257-7 (1-1161)

```
Q 4 MetAsnProGluLysArgTyrLeuPheLysLeuLeuLeuIleGlyAspSerGlyValGly 23
D 18 ATGGCAGAGACGATTAATTCCTTCAAGCTCTGCTGATCGGCGAGCTGGGGGATAGGC 77
Q 24 LysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyrIleSerThr 43
```

```
D 78 AAGACCTGCTCCGTGTCCTCCCTTCTCAGAGAGCGCTTCAACACCACCTTCATCTCCACC 137
Q 44 ILeGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLysLeuGln 63
D 138 ATCGAAATGTATTTTAAATTAAGACATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 197
Q 64 ILe----- 64
D 198 ATATGGGACACAGCGGGTCCAGAAAGATTCGAAACATCAAGACAGCGTACTACAGAGA 257
Q 65 -----GluSerPheAsnAsnValLys 71
D 258 GCCATGGGCAATTATGCTGCTGTATGACATCAAAATGAAAAATCTTGACAAATATTAA 317
Q 72 GlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuValGly 91
D 318 AATTGGATCAGAAACATTGAAAGCATGCCCTCTCCGATGTCGAAAGAAATGATCTGGGT 377
Q 92 AsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLysGluPheAla 111
D 378 AACAAATGTATGATGATGACAAAGAACAAAGTGTCAAAAGAAAGAGGAGGAGAACTAGCA 437
Q 112 AspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnValGluGln 131
D 438 ATTGACTATGGGATTAAATTCTTGAGACAAGCGCAAAATTCAGTGCAAATGTAGAAAGAG 497
Q 132 SerPheMetThrMetAlaIleGluIle-----LysLysArgMetGlyProGly 147
D 498 GCATTTTTCACCTTGACGACGAGATTAATGACAAACTCAACAGAAAAATGAATGACAGC 557
Q 148 AlaThrAlaGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGln 167
D 558 AATTACACGAGAGAGAGCT-----GACACAGTGAATAATACAGAAA-CCGATCAAGAA 610
```

RESULT 10

```
US-09-834-975-879
; Sequence 879, Application US/09834975
; Patent No. US2002010815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Hufel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MFI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 879
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-879
```

Alignment Scores:

```
Pred. No.: 1.51e-44 Length: 2497
Score: 414.00 Matches: 87
Percent Similarity: 60.508 Conservative: 34
Best Local Similarity: 43.508 Mismatches: 41
Query Match: 46.738 Indels: 39
10 Gaps: 3
```

US-09-820-003a-2 (1-173) x US-09-834-975-879 (1-2497)

OTHER INFORMATION: n = A, T, C or G
US-09-834-975-894

Alignment Scores:

Pred. No.:	1,51e+44	Length:	2497
Score:	414.00	Matches:	87
Percent Similarity:	60.50%	Conservative:	34
Best Local Similarity:	43.50%	Mismatches:	41
Query Match:	46.73%	Indels:	39
DB:	10	Gaps:	3

US-09-820-003A-2 (1-173) X US-09-834-975-894 (1-2497)

OY	4	MelnsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGly	23
Db	92	ATGGCGAAGACGTACGATATATCTTCAAGCTCCTGCTGATGGCGACTCGGGGGTGGC	155
OY	24	LysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyrIleSerThr	43
Db	152	AAGACCTCCTCTGTTCCGCTCTCAAGAGAGACGCCCTTCAACACACCTTCATCTCCACC	211
OY	44	IleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLysLeuGln	63
Db	212	ATCGGAATTTGATTTAAATTTAGAACGATACAGTACAGTGGAAAGAAATTAAGCTTCAG	277
OY	64	Ile	64
Db	272	ATATGGACACACGGCGGTACGAAAAGATTCCGAACATCACGACAGCGTACTACAGAGA	333
OY	65	-----GluSerPheAsnValLys	71
Db	332	GCCATGGGCAATTATGCTGCTGTATGATACATCAAAATGAAATATCTTGTGACAAATATTAA	392
OY	72	GlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeuLeuValGly	91
Db	392	AATTGGATCAAAACCATTTGAAGACATGCGCTTCGATGTCGAAAGATGATCTCGGCT	452
OY	92	AsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLysGluPheAla	111
Db	452	AACCAATGTGATATGATATGACAAAGAACACAGACTGTCCAAAGAGAAGAGGGGAGGACCTTGCA	511
OY	112	AspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsnValGluGln	131
Db	512	ATTGACTATGGGATTTAAATCTTGTGGAGACAGGCCAAATCCAGTGCAAATGTGAAAGAG	572
OY	132	SerPheMetThrMetAlaIleGluIle-----LysLysArgMetGlyProGly	147
Db	572	GCAATTTTTCACCTGTGCGAGAAATATATGACAAACTCAACAGAAATAATGAATGACAGC	633
OY	148	AlaThrIleArgLysValGluLysSerAsnValLysIleGlnSerThrProValGln	167
Db	632	AATTTCAGGAGGACGAGT-----GACCCATGTAAATATACAGAAA-CCGATTAAGGAA	684

RESULT 13
ITS-08-834-075-806

```

: Sequence 896, Application US/09834975
: Patent No. US20020110815A1
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Brown, Jeffrey
: APPLICANT: Bolt, Andrew
: APPLICANT: Van Bufler, Christophe
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
: TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
: FILE REFERENCE: MRI-016B
: CURRENT APPLICATION NUMBER: US/09/834,975
: CURRENT FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: 60/197,538
: PRIOR FILING DATE: 2000-04-14
: NUMBER OF SEQ ID NOS: 1046
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 896

```

```

: LENGTH: 2497
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(2497)
: OTHER INFORMATION: n = A,T,C or G
US-09-834-975-896

```

Alignment Scores:

Pred. No.:	1.51e-44	Length:	2497
Score:	414.00	Matches:	87
Percent Similarity:	60.50%	Conservative:	34
Best Local Similarity:	43.50%	Mismatches:	41
Query Match:	46.7%	Indels:	39
DB:	10	Gaps:	3

US-09-820-003A-2 (1-173) x US-09-834-975-896 (1-2497)

RESULT 14
HS-00-02E-300-631

: Sequence 631, Application US/09925300
 : Patent No. US20070151681A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Craig Rosen,
 : APPLICANT: Steve Ruben
 : TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 : FILE REFERENCE: PA101
 : CURRENT APPLICATION NUMBER: US/09/925,300
 : CURRENT FILING DATE: 2001-08-10
 : PRIOR APPLICATION NUMBER: PCT/US00/05988
 : PRIOR FILING DATE: 2000-03-08
 : PRIOR APPLICATION NUMBER: 60/124,270

Search completed: November 17, 2002, 14:31:59
Job time : 62 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 13:01:57 ; Search time 43 Seconds
(without alignments)
386,773 Million cell updates/sec

Title: us-09-820-003a-2

Perfect score: 886
Sequence: 1 MSSMNPEDYLFKLLIGDS.....EKSNNKIQSTPVKQSGGGCC 173

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	860	97.1	205	1 TVHUTP	GTP-binding protein
2	860	97.1	205	1 TVDSTP	GTP-binding protein
3	860	97.1	205	1 TVKSTP	GTP-binding protein
4	849	95.8	205	1 TVRSTP	GTP-binding protein
5	812.5	91.7	201	2 D38625	GTP-binding protein
6	771.5	87.1	201	2 S06147	GTP-binding protein
7	733	82.7	205	2 S38339	GTP-binding protein
8	671	75.7	202	2 JE0318	GTP-binding protein
9	660	74.5	205	2 T33781	hypothetical protein
10	658.5	74.3	203	2 S30096	GTP-binding protein
11	646.5	73.0	203	2 JC4105	GTP-binding protein
12	646.5	73.0	203	2 T50323	GTP-binding protein
13	642.5	72.5	203	2 JC1247	GTP-binding protein
14	631.5	71.3	201	2 JC5337	GTP-binding protein
15	627.5	70.8	203	2 S34253	GTP-binding protein
16	623.5	70.4	206	2 S04590	GTP-binding protein
17	619.5	69.9	258	2 B86153	GTP-binding protein
18	619	69.9	202	2 S41430	GTP-binding protein
19	618.5	69.8	203	2 B38202	GTP-binding protein
20	613.5	69.1	221	2 S38740	GTP-binding protein
21	612	69.1	221	2 H71444	GTP-binding protein
22	598	67.5	202	2 T07609	GTP-binding protein
23	590.5	66.6	202	2 S72515	GTP-binding protein
24	587	66.3	201	2 S39565	GTP-binding protein
25	574	64.8	206	2 T14391	GTP-binding protein
26	572.5	64.6	196	2 PS0279	GTP-binding protein
27	571	64.4	206	1 TVR02	GTP-binding protein
28	523	59.0	208	2 A38202	GTP-binding protein
29	468.5	52.9	208	2 A34716	GTP-binding protein

30	461	52.0	203	2 B34716	GTP-binding protein
31	447	50.5	216	2 T48378	GTP-binding protein
32	442	49.9	215	2 T14565	GTP-binding protein
33	439.5	49.6	215	2 S57478	GTP-binding protein
34	438.5	49.5	206	2 J50640	GTP-binding protein
35	435.5	49.2	215	2 S12790	GTP-binding protein
36	430.5	48.6	215	2 S57462	GTP-binding protein
37	429.5	48.5	215	2 S57471	GTP-binding protein
38	429.5	48.5	222	2 T14405	GTP-binding protein
39	428.5	48.4	216	2 T45901	GTP-binding protein
40	426.5	48.1	215	2 S57474	GTP-binding protein
41	425.5	48.0	216	2 S33900	GTP-binding protein
42	418	47.2	207	2 B49647	GTP-binding protein
43	418	47.2	207	2 B36364	GTP-binding protein
44	410.5	46.3	206	2 T17851	GTP-binding protein
45	409.5	46.2	209	2 B38625	GTP-binding protein

ALIGNMENTS

RESULT 1

GTP-binding protein Rabi - human

N:Alternate names: protein DKFZp564B163.1

C:Species: Homo sapiens (man)

C>Date: 29-Jun-1990 #sequence_revision 06-Dec-1996 #text_change 19-Jan-2001

C:Accession: A34323; T08698

R:Zahroul, A.; Touchot, N.; Chardin, P.; Tavallan, A.
J. Biol. Chem. 264, 12394-12401, 1989

A:Title: The human Rabi genes encode a family of GTP-binding proteins related to yeast.

A:Reference number: A34323; MUID:89308668; PMID:2501306

A:Accession: A34323

A:Molecule type: mRNA

A:Residues: 1-205 <ZMW>

A:Cross-references: GB:04941; GB:M28209; NID:9550059; PIDN:AAA60240.1; PID:9550060

A:Experimental source: EMBL:AL050268

A:Gene: GDB:RAB1

A:Cross-references: GDB:118857; OMIM:179508

A:Map position: 4P15.31-4P15.31

A>Note: DKFZp564B163.1

C:Function: probably involved in protein transport from the endoplasmic reticulum

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipid protein; membran

tein

F:1-205/Product: GTP-binding protein Rabi1 #status predicted <NNT>

F:1-64,141-205/Product: GTP-binding protein Rabi1, splice variant #status predicted <K

F:12-127/Domain: translation elongation factor Tu homology <ETU>

F:18-25/Region: nucleotide-binding motif A (P-loop)

F:62-67/Region: nucleotide-binding motif B

F:124-127/Region: GTP-binding NKXD motif

F:154-156/Region: GTP-binding SAFL motif

F:24,25,43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #

F:124/Binding site: phosphate (Ser) (covalent) #status predicted

F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status experimental

Query Match 97.1% Score 860; DB 1; Length 205;

Best Local Similarity 84.4% Pred. No. 2,2e+62;

Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

Qy 1 MSSMNPEDYLFKLLIGDSGVKSCLLRFADDTYTESYSTGDFKRIELDKGTI 60

Db 1 MSSMNPEDYLFKLLIGDSGVKSCLLRFADDTYTESYSTGDFKRIELDKGTI 60


```
C:Function:
A:Description: probably involved in protein transport from the endoplasmic reticulum th
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: endoplasmic reticulum; Golgi apparatus; GTP binding; lipoprotein; membrane p
tein
F:12-127/Domain: translation elongation factor Tu homology <ETU>
F:18-25/Region: nucleotide-binding motif A (P-loop)
F:62-67/Region: nucleotide-binding motif B
F:124-127/Region: GTP-binding NKX motif
F:154-156/Region: GTP-binding SAK/L motif
F:24/25/43,124,125,127,154/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
F:194/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status predicted
F:204,205/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match          95.8%; Score 849; DB 1; Length 205;
Best Local Similarity 83.4%; Pred. No. 1,7e-61;
Matches 171; Conservative 0; Mismatches 2; Indels 32; Gaps 1;

OY 1 MSNMPEYDLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDFKIRTELDGKTI 60
    |||||||
Db 1 MSNMPEYDLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDFKIRTELDGKTI 60
OY 61 KIQI-----ESFNNVKOMLOEIDRYASBNKLV 88
    |||||||
Db 61 KIQIDPAGGERFRTTSSYRGAHGIYVDVDTQESFNNVKOMLOEIDRYASBNKLV 120
OY 89 LVGNKCDLTKKVVDTTAKFEADSLGIPLETSAKNATNVEOSFMTMAEIKRMGPGA 148
    |||||||
Db 121 LVGNKCDLTKKVVDTTAKFEADSLGIPLETSAKNATNVEOSFMTMAEIKRMGPGA 180
OY 149 TAGAENKSNKIOSTPVKOSGGGCC 173
    |||||||
Db 181 TAGAENKSNKIOSTPVKOSGGGCC 205

RESULT 5
GTP-binding protein o-rab1 - electric ray (Discopyge ommata)
C:Species: Discopyge ommata
C>Date: 23-Aug-1991 #sequence_rev150n 23-Aug-1991 #text_change 02-Feb-2001
C:Accession: D38625
R:Ngase, U.K.; Eitenink, L.A.; Scheller, R.H.
J: Biol. Chem. 266, 2675-2680, 1991
A:Title: A family of ras-like GTP-binding proteins expressed in electromotor neurons.
A:Reference number: A38625; MUID:91115900; PMID:1899244
A:Accession: D38625
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-201 <NGS>
A:Cross-references: GB:M38393
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NKX motif
F:150-152/Region: GTP-binding SAK/L motif
F:21,22,40,121,122,124,150/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
F:200,201/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match          91.7%; Score 812.5; DB 2; Length 201;
Best Local Similarity 81.2%; Pred. No. 1.5e-58;
Matches 164; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

OY 4 MNPEYDLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDFKIRTELDGKTI 63
    |||||||
Db 1 MNPEYDLFKLLIDSGVSKCLLRFPADDTYTESYSTIGVDFKIRTELDGKTI 60
OY 64 I-----ESFNNVKOMLOEIDRYASBNKLV 91
    |||||||
Db 61 IMDPAGGERFRTTSSYRGAHGIYVDVDTQESFNNVKOMLOEIDRYASBNKLV 120
OY 92 NKCDLTKKVVDTTAKFEADSLGIPLETSAKNATNVEOSFMTMAEIKRMGPGA 151
    |||||||
Db 121 NKSDLTTKKVVDTTAKFEADSLGIPLETSAKNATNVEOSFMTMAEIKRMGPGA 180
OY 152 GAENKSNKIOSTPVKOSGGGCC 173
    |||||||
Db 181 G-ERNLKLIDSTPVKASGGCC 201

RESULT 7
GTP-binding protein rab1 - great pond snail
C:Species: Lymnaea stagnalis (great pond snail)
C>Date: 13-Jan-1995 #sequence_rev150n 13-Jan-1995 #text_change 02-Feb-2001
C:Accession: S38339; S32206
R:Agteberg, M.; van Die, I.; Yang, H.; Andriessen, J.A.; van Teerling, A.; van den E
Bur, J. Biochem. 217, 241-246, 1993
A:Title: Isolation and characterization of three cDNAs coding for Rab proteins from t
A:Reference number: S38339; MUID:94039042; PMID:8223561
A:Accession: S38339
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 11:42:42 : Search time 25 Seconds
(without alignments)
287.016 Million cell updates/sec

Title: US-09-820-003a-2
Perfect score: 886
Sequence: 1 MSNMPEYDYLFKLLIGDS.....EKSNNKISTEVKSGGCC 173

Scoring table: BL0SUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	860	97.1	205	1 RB1A_HUMAN	P11476 homo sapien
2	849	95.8	205	1 RB1A_HUMAN	P05711 ratius norv
3	828	93.5	202	1 RB1A_HUMAN	P22123 discopige o
4	760.5	85.8	201	1 RB1A_HUMAN	P10536 ratius norv
5	732	82.7	203	1 RB1A_HUMAN	P03974 lymnaea sta
6	658.2	74.3	203	1 RB1A_HUMAN	P33723 neuropept
7	646.5	73.0	203	1 RB1A_HUMAN	P33571 chlamydomon
8	646.5	73.0	203	1 RB1A_HUMAN	P11620 schizosach
9	642.5	72.5	203	1 RB1A_HUMAN	P31584 volvox cart
10	631.5	71.3	201	1 RB1A_HUMAN	P01890 phytophthor
11	619.5	69.9	258	1 RB1A_HUMAN	P01890 phytophthor
12	618.5	69.8	203	1 RB1A_HUMAN	P28188 arabidopsis
13	613.5	69.2	202	1 RB1A_HUMAN	P00577 zea mays (m
14	571	64.4	206	1 RB1A_HUMAN	P40392 oryza sativ
15	542	61.2	199	1 RB1A_HUMAN	P01133 saccharomyc
16	538	60.7	199	1 RB1A_HUMAN	P34139 dictyostell
17	523	59.0	208	1 RB1A_HUMAN	P34140 dictyostell
18	468.5	52.9	208	1 RB1A_HUMAN	P16976 zea mays (m
19	461	52.0	203	1 RB1A_HUMAN	P20790 dictyostell
20	442	49.9	215	1 RB1A_HUMAN	P20791 dictyostell
21	438.5	49.5	216	1 RB1A_HUMAN	P03433 beta vulgar
22	435.5	49.2	200	1 RB1A_HUMAN	P28186 arabidopsis
23	424	47.9	210	1 RB1A_HUMAN	P17609 schizosach
24	418	47.2	207	1 RB1A_HUMAN	P22128 discopige o
25	417	46.8	207	1 RB1A_HUMAN	P24407 homo sapien
26	414.5	46.8	199	1 RB1A_HUMAN	P70550 ratius norv
27	414	46.3	207	1 RB1A_HUMAN	P31411 dictyostell
28	410.5	45.8	206	1 RB1A_HUMAN	P02930 homo sapien
29	405.5	45.5	217	1 RB1A_HUMAN	P55258 mus musculu
30	403.5	45.8	200	1 RB1A_HUMAN	P36861 volvox cart
31	402.5	45.4	200	1 RB1A_HUMAN	P08366 homo sapien
32	398	44.9	200	1 RB1A_HUMAN	P24409 canis famli
33	398	44.9	203	1 RB1A_HUMAN	P21217 discopige o
					P41924 yarrowia li

34	384.5	43.4	203	1 RB13_HUMAN	P51153 homo sapien
35	374.5	42.3	200	1 RB10_RAT	P35281 ratius norv
36	373.5	42.2	215	1 RB10_RAT	P07560 saccharomyc
37	360.5	40.7	201	1 RB35_HUMAN	O15286 homo sapien
38	360.5	40.6	210	1 RB35_HUMAN	O14462 candida alb
39	352	39.7	209	1 RB2A_MAIZE	P49103 zea mays (m
40	349.5	38.9	210	1 RB2A_MAIZE	P49104 zea mays (m
41	344.5	38.9	213	1 YPT4_CHLRE	P03970 chlamydomon
42	342	38.6	219	1 RB3A_BOVIN	P11023 bos taurus
43	339	38.3	213	1 YPT4_YOLCA	P36863 volvox cart
44	336	37.9	220	1 RBAB3_DROME	P25228 drosophila
45	336	37.9	220	1 RB3A_HUMAN	P20336 homo sapien

ALIGNMENTS

RESULT 1
ID RB1A_HUMAN STANDARD; PRT; 205 AA.
AC P11476;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-1A (YPT1-related protein).
GN RAB1 OR RAB1A.
OS Homo sapiens (Human), and
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606, 10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; PubMed=2501306.
RX MEDLINE=89308668; PubMed=2501306.
RA Zahraoui A., Touchot N., Chardin P., Tavittian A.;
RT "The human Rab genes encode a family of GTP-binding proteins related
to yeast YPT1 and SEC4 products involved in secretion.";
J Biol. Chem. 264:12394-12401(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; Trisub-Placenta;
RX Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE=88166649; PubMed=3127202;
RA Haubruck H., Disela C., Wagner P., Gallwitz D.;
RT "The ras-related ypt protein is an ubiquitous eukaryotic protein:
isolation and sequence analysis of mouse cDNA clones highly
homologous to the yeast YPT1 gene.";
EMBO J. 6:4049-4053(1987).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE=89386011; PubMed=2506528;
RA Wichmann H., Disela C., Haubruck H., Gallwitz D.;
RT "Nucleotide sequence of the mouse ypt1 gene encoding a ras-related
GTP-binding protein.";
Nucleic Acids Res. 17:6737-6738(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX W.G., Dorn G.W., II;
RA "Mouse Rab1a, member of Ras oncogene family mRNA.";
Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP PHOSPHORYLATION BY CDC2.
RX MEDLINE=9121885; PubMed=1902553;
RA Bailey E., McCreedy M., Touchot N., Zahraoui A., Goud B., Bornens M.;
RT "Phosphorylation of two small GTP-binding proteins of the Rab family
by p34cdc2.";

```

RL Nature 350:715-718(1991).
RN [7]
RP ISOPRENOLID.
RX MEDLINE-91296801; Pubmed=1648736;
RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,
RT "Isoprenoid modification of rab proteins terminating in CC or CXC
molts."
RL Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268(1991).
CC -1- FUNCTION: PROBABLY REQUIRED FOR TRANSIT OF PROTEIN FROM THE ER
CC THROUGH GOLGI COMPARTMENT.
CC -1- SUBCELLULAR LOCATION: GOLGI COMPARTMENT.
CC -1- PPM: PHOSPHORYLATED BY CDC2-KINASE DURING MITOSIS.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M28209; AAA60240.1; -.
DR EMBL: BC000905; AAR00905.1; -.
DR EMBL: Y00094; CAA68284.1; -.
DR EMBL: X15744; CAA33760.1; -.
DR EMBL: X15745; CAA33760.1; JOINED.
DR EMBL: X15746; CAA33760.1; JOINED.
DR EMBL: X15747; CAA33760.1; JOINED.
DR EMBL: AF226873; AAF33844.1; -.
DR PIR: S05551; TVMSIP.
DR PIR: A34323; A34323.
DR PIR: S19104; S19104.
DR HSSP: P05713; 3RAB.
DR PMAA-2DPAGE: P11476; -.
DR Genew: HNCN:9758; RAB1A.
DR MIM: 179508; -.
DR MGD: MGI:97842; Rab1.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFS: TIGR00231; small_gtp; 1.
DR GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack;
KW Endoplasmic reticulum; Phosphorylation.
KM NP_BIND 18 25 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 124 127 GTP (BY SIMILARITY).
FT DOMAIN 40 48 EFECTOR REGION (BY SIMILARITY).
FT MOD_RES 194 194 PHOSPHORYLATION (BY CDC2) (PROBABLE).
FT LIPID 204 204 GERANYL-GERANYL.
FT LIPID 205 205 GERANYL-GERANYL.
SQ SEQUENCE 205 AA; 22678 MW; B2ABF4E3B0FB1D6 CRC64;
Query Match 97.1%; Score 860; DB 1; Length 205;
Best local Similarity 84.4%; Pred. No. 9.3e-65;
Matches 173; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

```

```

QY 149 TAGAERKSNVKIOSTPVKOSGGGCC 173
DB 181 TAGAERKSNVKIOSTPVKOSGGGCC 205
RESULT 2
ID RBA1_RAT STANDARD; PRT; 205 AA.
AC P05711;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ras-related protein Rab-1A.
GN RAB1 OR RAB1A.
OS Rattus norvegicus (Rat), and
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116, 9615;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Rat;
RX MEDLINE-88068563; Pubmed=3317403;
RA Touchot N., Chardin P., Tavittian A.;
RT "Four additional members of the ras gene superfamily isolated by an
RT oligonucleotide strategy: molecular cloning of YPT-related cDNAs from
RT a rat brain library."
RL Proc. Natl. Acad. Sci. U.S.A. 84:8210-8214(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-C.familiaris; STRAIN-Cocker spaniel;
RX MEDLINE-91061765; Pubmed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
RT line."
RL Mol. Cell. Biol. 10:6578-6585(1990).
RN [3]
RP ISOPRENOLID.
RX MEDLINE-91296801; Pubmed=1648736;
RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,
RA Shensky M., Balch W.E., Buss J.E., Der C.J.;
RT "Isoprenoid modification of rab proteins terminating in CC or CXC
molts."
RL Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268(1991).
CC -1- FUNCTION: PROBABLY REQUIRED FOR TRANSIT OF PROTEIN FROM THE ER
CC THROUGH GOLGI COMPARTMENT.
CC -1- MISCELLANEOUS: RAB-1A BINDS GTP AND GDP AND POSSESS INTRINSIC
CC GTPASE ACTIVITY.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02998; AAA42006.1; -.
DR EMBL: X56384; CAB56775.1; ALT-SEQ.
DR PIR: A36364; A36364.
DR PIR: A39963; A39963.
DR HSSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFS: TIGR00231; small_gtp; 1.
DR GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack;
KW Endoplasmic reticulum.
FT NP_BIND 18 25 GTP (BY SIMILARITY).

```

```

FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 124 127 GTP (BY SIMILARITY).
FT DOMAIN 40 48 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL.
FT LIPID 205 205 GERANYL-GERANYL.
SQ SEQUENCE 205 AA: 22763 MW: 34649523BDF0DEB CRC64:

Query Match 95.8%; Score 849; DB 1; Length 205;
Best Local Similarity 83.4%; Pred. No. 7.7e-64;
Matches 171; Conservative 0; Mismatches 2; Indels 32; Gaps 1;

QY 1 MSNMPEDYFLKLLIGDSGVKSCLLRPADDTYESTISTGVDFKRTIELDKGTI 60
   |||||
DB 1 MSNMPEDYFLKLLIGDSGVKSCLLRPADDTYESTISTGVDFKRTIELDKGTI 60
   |||||
QY 61 KIQT-----ESFNNVKOMLOEIDRVASENVKLVG 88
   |||||
DB 61 KIQTDRAGERRRTTSSYRGAGIIVYDVYDOESFNNVKOMLOEIDRVASENVKLVG 120
   |||||
QY 89 LVGNKCDLTTRKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRRMGPGA 148
   |||||
DB 121 LVGNKCDLTTRKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRRMGPGA 180
   |||||
QY 149 TAGAEKSNVKIQSTPVKQSGGCC 173
   |||||
DB 181 TAGAEKSNVKIQSTPVKQSGGCC 205

RESULT 3
RAB1_DISOM STANDARD; PRT; 202 AA.
AC P22125;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Ras-related protein ORAB-1.
OS Discopyle ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hypnosqualea; Pristigaster; Batoidae;
OC Torpediniformes; Narcinidae; Narciniidae; Discopyle.
OX NCBI_Taxid=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Electric lobe;
RX MEDLINE=91115900; PubMed=1899244;
RA Ngeese J.K., Elferink L.A., Scheller R.H.;
RT "A family of ras-like GTP-binding proteins expressed in electromotor
   neurons."
RT J. Biol. Chem. 266:2675-2680(1991).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
   TRAFFIC (BY SIMILARITY).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL Outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (see http://www.isb-sdb.ch/announce/
   or send an email to license@isb-sdb.ch).
CC EMBL: M38393; AAA9234.1; -.
DR FHS: D38625; D38625.
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_Transfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras_1.
DR PRINTS: PR00449; RASTRNSFRNG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_gtp; 1.
KM GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).

```

```

FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 202 AA: 22333 MW: 31078E502BCDD1B9 CRC64:

Query Match 93.5%; Score 828; DB 1; Length 202;
Best Local Similarity 81.7%; Pred. No. 4.2e-62;
Matches 165; Conservative 3; Mismatches 2; Indels 32; Gaps 1;

QY 4 MNPEDYFLKLLIGDSGVKSCLLRPADDTYESTISTGVDFKRTIELDKGTIKQ 63
   |||||
DB 1 MNPEDYFLKLLIGDSGVKSCLLRPADDTYESTISTGVDFKRTIELDKGTIKQ 60
   |||||
QY 64 I-----ESFNNVKOMLOEIDRVASENVKLVG 91
   |||||
DB 61 IWDPAQERFRTTSSYRGAGIIVYDVYDOESFNNVKOMLOEIDRVASENVKLVG 120
   |||||
QY 92 NKCDLTTRKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRRMGPGAAG 151
   |||||
DB 121 NKCDLTTRKVVDTTAKFADSLGIPLETSAKNATNVEOSFMTMAEIKRRMGPGAAG 180
   |||||
QY 152 GAESKSNVKIQSTPVKQSGGCC 173
   |||||
DB 181 GSEKSNVKIQSTPVKQSGGCC 202

RESULT 4
RAB1B_RAT STANDARD; PRT; 201 AA.
AC P10536;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Ras-related protein Rab-1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89160341; PubMed=2493636;
RA Zahraoui A., Touchot N., Charotin P., Tavittian A.;
RT "Nucleotide sequence of a rat cDNA: rab1b, encoding a rab1-YPT
   related protein."
RT Nucleic Acids Res. 17:1770-1770(1989).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=9003316; PubMed=2509243;
RA Touchot N., Zahraoui A., Viehn E., Tavittian A.;
RT "Biochemical properties of the YPT-related rab1b protein. Comparison
   with rab1A."
RT FEBS Lett. 256:79-84(1989).
RN [3]
RP ISOPRENOLID.
RX MEDLINE=91296801; PubMed=1648736;
RA Khosravi-Far R., Lutz R.J., Cox A.D., Conroy L., Bourne J.R.,
RA Sienksky M., Balch W.E., Buss J.E., Der C.J.;
RT "Isoprenoid modification of rab proteins terminating in CC or CXC
   motifs."
RT Proc. Natl. Acad. Sci. U.S.A. 88:6264-6268(1991).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
   TRAFFIC (BY SIMILARITY).
CC -1- MISCELLANEOUS: RAB-1B BINDS GTP AND GDP AND POSSESS INTRINSIC
   GTPASE ACTIVITY.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL Outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X13905; CAA32105.1; -.
DR PIR: S03189; S03189.
DR PIR: S06147; S06147.
DR HSSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMNC.
DR SMART: SM00175; RAB; 1.
DR TIGRFRAMS: TIGR00231; small_gtp; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NE_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT DOMAIN 121 124 GTP (BY SIMILARITY).
FT LIPID 200 200 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL.
FT MUTAGEN 21 21 K->M: ABOLISHES GTP-BINDING.
SQ SEQUENCE 201 AA; 22163 MW; 8D3EDC2AEF4A2PE CRC64;

Query Match 85.8%; Score 760.5; DB 1; Length 201;
Best Local Similarity 76.2%; Pred. No. 1,7e-56;
Matches 154; Conservative 7; Mismatches 8; Indels 33; Gaps 2;

QY 4 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTGVDFKRTIELDKTI 63
DB 1 MNPEYDLFKLLIGDSGVGKSCLLRFADDTYESYISTGVDFKRTIELDKTI 60
QY 64 I-----ESFNNVKOMLOEIDRYASENVNKL 91
DB 61 IMDPAGGERRTYSSYRGAGHGIIVYDVTDOESYANVKOMLOEIDRYASENVNKL 120
QY 92 NKCDLTKRKVVYDTTAKFADSLGIPLETSAKNATVEQSFMTMAEIKRMPGATAG 151
DB 121 NKSDLTKRKVVYDTTAKFADSLGIPLETSAKNATVEQSFMTMAEIKRMPGATAG 180
QY 152 GAESKNVKIOSTPYKOSGGCC 173
DB 181 G-ERPNIKIDSTPYKASAGGCC 201

RESULT 5
RABL_LYNSF
ID RABL_LYNSF STANDARD; PRT; 205 AA.
AC 005974;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Ras-related protein Rab-1A.
GN RAB1A.
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94039042; PubMed=8223561;
RA Agterberg M., van Die I., Yang H., Andriessen J.A.,
van Tetering A., van den Elinden D.H., Ploegh H.L.;
RT "Isolation and characterization of three cDNAs coding for Rab
proteins from the alveolar gland of the mollusc Lymnaea stagnalis.";
RL Eur. J. Biochem. 217:241-246(1993).
CC -1- FUNCTION: PROBABLY REQUIRED FOR TRANSIT OF PROTEIN FROM THE ER
THROUGH GOLGI COMPARTMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: GOLGI COMPARTMENT (BY SIMILARITY).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUPERFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X72688; CAA51233.1; -.
DR PIR: S32206; S32206.
DR PIR: S38339; S38339.
DR HSSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMNC.
DR SMART: SM00175; RAB; 1.
DR TIGRFRAMS: TIGR00231; small_gtp; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack;
KW Endoplasmic reticulum.
FT NE_BIND 18 25 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 124 127 GTP (BY SIMILARITY).
FT DOMAIN 40 48 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 205 AA; 22760 MW; D5E0BDDCB84A26FF CRC64;

Query Match 82.7%; Score 733; DB 1; Length 205;
Best Local Similarity 73.8%; Pred. No. 3.4e-54;
Matches 152; Conservative 7; Mismatches 13; Indels 34; Gaps 3;

QY 1 MSSNPEDYLFKLLIGDSGVGKSCLLRFADDTYESYISTGVDFKRTIELDKTI 60
DB 1 MSTNMPDYLFKLLIGDSGVGKSCLLRFADDTYESYISTGVDFKRTIELDKTI 60
QY 61 KLOI-----ESFNNVKOMLOEIDRYASENVNKL 88
DB 61 KLOI-----ESFNNVKOMLOEIDRYASENVNKL 120
QY 89 LVGNKCDLTKRKVVYDTTAKFADSLGIPLETSAKNATVEQSFMTMAEIKRMPGA 148
DB 121 LVGNKCDLTKRKVVYDTTAKFADSLGIPLETSAKNATVEQSFMTMAEIKRMPGA 179
QY 149 TAGGAESKNVKI-OSTPYKOSGGCC 173
DB 180 TAASDSKPSVKINSSTPYKANKGCC 205

RESULT 6
YPTL_NEUCR
ID YPTL_NEUCR STANDARD; PRT; 203 AA.
AC P33723;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein ypt1.
GN YPT-1 OR B9J10.240.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93101148; PubMed=1361212;
RA Heintz K., Palme K., Diefenthal T., Russo V.E.A.;
RT "The Ncylt1 gene from Neurospora crassa is located on chromosome 2:
RT molecular cloning and structural analysis.";
RL Mol. Genet. 235:413-421(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RA Schulte U., Algen V., Hohelsel J., Brandt P., Fartmann B., Holland R.,

```

RA Nyakatura G., Mewes H.-W., Mannhaupt G.;
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S51252; AAB24564.1; -
DR EMBL: AL356324; CAB92031.1; -
DR PIR: S30096; S30096.
DR HSSP: F01112; 1PLU.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_transfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRNG.
DR SMART: SM00175; RAB; 1.
DR TIGRfam: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
KM NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT LIPID 202 202 GERANYL-GERANYL.
FT LIPID 203 203 GERANYL-GERANYL.
SQ SEQUENCE 203 AA; 22476 MW; 20F5A1AB3400EFD CRC64;

Query Match 74.3%; Score 658.5; DB 1; Length 203;
Best Local Similarity 66.5%; Pred. No. 5.3e-48;
Matches 135; Conservative 11; Mismatches 24; Indels 33; Gaps 2;

OY 4 MNPEVDYLFKLLIGDSGVKSCLLLRPADDTYESYSTIGVDFKIRTELDGKTIKQ 63
DB 1 MNPEVDYLFKLLIGDSGVKSCLLLRPADDTYESYSTIGVDFKIRTELDGKTIKQ 60
OY 64 I-----ESFNWVKOMLQEIIDRYASENVKLLVG 91
DB 61 IWDVAGQERFRTITSSYRGAGHICVYDVTDMDSEFNWVKOMLQEIIDRYATEGVNKLIVG 120
OY 92 NKKDILTKKVVDTTAKERADSLGIPLETSAKNAATNVEOSFMTMAEIKRKG--PCAT 149
DB 121 NKSDDTKKVVETTAKERADSLGIPLETSAKNAATNVEOAFETMAROKERKSSIAATN 180
OY 152 GAERK-S-NVKIGSTPVKQSGGCC 173
DB 181 NTKASVNVSPGHCVSNNSSGCC 203

RESULT 7
YPTL_CHLRE
ID YPTL_CHLRE STANDARD; PRT; 203 AA.
AC Q39571.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GTP-binding protein YPT1.
GN YPT1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadales; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CA15;
RX MEDLINE=95309723; PubMed=7789809;
RA Dietmaier W., Fabry S., Huber H., Schmitt R.;

```

```

FT "Analysis of a family of ypt genes and their products from
RL Chlamydomonas reinhardtii.",
RL Gene 158:41-50(1995).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U13168; AAA82727.1; -
DR HSSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_transfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRNG.
DR SMART: SM00175; RAB; 1.
DR TIGRfam: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;
KM Multigene family.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 203 AA; 22599 MW; DD24240F2F35EID CRC64;

Query Match 73.0%; Score 646.5; DB 1; Length 203;
Best Local Similarity 65.7%; Pred. No. 5.3e-47;
Matches 144; Conservative 9; Mismatches 26; Indels 35; Gaps 3;

OY 4 MNPEVDYLFKLLIGDSGVKSCLLLRPADDTYESYSTIGVDFKIRTELDGKTIKQ 63
DB 1 MNPEVDYLFKLLIGDSGVKSCLLLRPADDTYESYSTIGVDFKIRTELDGKTIKQ 60
OY 64 I-----ESFNWVKOMLQEIIDRYASENVKLLVG 91
DB 61 IWDVAGQERFRTITSSYRGAGHICVYDVTDMDSEFNWVKOMLQEIIDRYASENVKLLVG 120
OY 92 NKKDILTKKVVDTTAKERADSLGIPLETSAKNAATNVEOSFMTMAEIKRKG--PCAT 149
DB 121 NKSDDTKKVVETTAKERADSLGIPLETSAKNAATNVEOAFETMAROKERKSSIAATN 180
OY 150 AGGAERKSNVKIGSTPVKQSGGCC 173
DB 181 KAGGPVVRPO--EGKPIKSSKSSCC 203

RESULT 8
YPTL_SCHPO
ID YPTL_SCHPO STANDARD; PRT; 203 AA.
AC P11620.
DT 01-OCT-1988 (Rel. 12, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein ypt1.
GN YPT1 OR SPAC1703.10.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=90228338; PubMed=2328721;
RX MEDLINE=90228338; PubMed=2328721;
RA Miyake S., Yamamoto M.;

```

RT "Identification of ras-related, YPT family genes in
 RT Schizosaccharomyces pombe.";
 RL EMBL J. 9:1417-1422(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-69236485; PubMed-2740223;
 RA Fawell E., Hook S., Armstrong J.;
 RT "Nucleotide sequence of a gene encoding a YPT1-related protein from
 RT Schizosaccharomyces pombe.";
 RL Nucleic Acids Res. 17:4373-4373(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RX MEDLINE-21848401; PubMed-11859360;
 RA Wood V., Gwilliam R., Rajendram M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Snelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakeert G., Aert R., Robben J., Grymopre B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabriel C., Fuchs M., Filiz C., Holzer E., Moestl D., Hübner H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hunt S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Gerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shipakowski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 RN [4]
 RP ISOPRENOLD.
 RX MEDLINE-92283842; PubMed-1597466;
 RA Newman C.M., Giannakouris T., Hancock J.F., Fawell E.H., Armstrong J.,
 RA Magee A.I.;
 RT "Post-translational processing of Schizosaccharomyces pombe YPT
 RT proteins.";
 RL J. Biol. Chem. 267:11329-11336(1992).
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X52099; CAA36319.1; -;
 CC EMBL: X15082; CAA3192.1; -;
 CC EMBL: AL136536; CAB66454.1; -;
 CC PIR: S04590; S04590.
 CC PIR: S10025; S10025.
 CC HSP: P05713; 3RAB.
 CC InterPro: IPR003579; GTPase_Rab.
 CC InterPro: IPR001806; Ras_trnsfmg.
 CC InterPro: IPR005225; Small_GTP.
 CC Pfam: PF00071; ras.1.
 CC PRINTS: PR00449; RASTRNSFRMG.
 CC SMART: SM00175; RAB; 1.
 CC TIGRfams: TIGR00231; small_gtp; 1.

KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
 FT NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 121 124 GTP (BY SIMILARITY).
 FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
 FT LIPID 202 202 GERANYL-GERANYL.
 FT LIPID 203 203 GERANYL-GERANYL.
 FT CONFLICT 1 4 MNPE -> NMANINR (IN REF. 2).
 FT SEQUENCE 203 AA; 22816 MW; 59B3524F42B3094 CRC64;
 Query Match 73.08; Score 646.5; DB 1; Length 203;
 Best Local Similarity 64.38; Pred. No. 5.3e-47;
 Matches 131; Conservative 16; Mismatches 23; Indels 33; Gaps 2;
 QY 4 MNPEYDLFLKLLIGDGVGKSCLLRFADDTYESYISTIGVDFKRTLELDTKTIKQ 63
 DB 1 MNPEYDLFLKLLIGDGVGKSCLLRFADDTYESYISTIGVDFKRTLELDTKTIKQ 60
 QY 64 I-----ESFNWKQWLOEIDRYASENVKLLVG 91
 DB 61 IWDYAGEFRFTTSSYRGAGIIVDYTDQDSFNWVKWQLOEIDRYAVEGVNRLVG 120
 QY 92 NKCLTLTKRYVDYTAKEPADSLGIFPLETSAKATNVEOSFMTMAEIKRMGPATAG 151
 DB 121 NKSDMVKRKYEVYSAKEFADSLNIPPLETSAKSTNVEQAFITMSHOIKERMGNTFAS 180
 QY 152 GAERKNYKI-QSTPVKOSGGCC 173
 DB 181 SNAKSSVKVGGGTNVSOSSMCC 203
 RESULT 9
 YPT1 VOLCA STANDARD; PRT; 203 AA.
 ID YPT1 VOLCA
 AC P31584;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GTP-binding protein ypt1.
 GN YPT1.
 OS Volvox carterl.
 CC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 CC Volvocaceae; Volvox.
 CC NCBI_TaxID=3067;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-f. Nagariensis / HK10;
 RX MEDLINE-92380499; PubMed-1511889;
 RA Fabry S., Nass N., Huber H., Palme K., Jaenicke L., Schmitt R.;
 RT "The ypt1 gene encodes a small G-protein in the green alga Volvox
 RT carterl: gene structure and properties of the gene product.";
 RL Gene 118:153-162(1992).
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M93438; AAA34255.1; -;
 CC PIR: JCI247; JCI247.
 CC HSP: P05713; 3RAB.
 CC InterPro: IPR003579; GTPase_Rab.
 CC InterPro: IPR001806; Ras_trnsfmg.
 CC InterPro: IPR005225; Small_GTP.
 CC Pfam: PF00071; ras.1.
 CC PRINTS: PR00449; RASTRNSFRMG.
 CC SMART: SM00175; RAB; 1.

```

DR  TIGR00231: small_GTP; 1.
KW  GTP-binding; Lipoprotein; Prenylation; Protein transport;
KM  Multigene family.
FT  NP_BIND 15 22 GTP (BY SIMILARITY).
FT  NP_BIND 63 67 GTP (BY SIMILARITY).
FT  NP_BIND 121 124 GTP (BY SIMILARITY).
FT  DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT  LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).
FT  LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).
SQ  SEQUENCE 203 AA: 22532 MW: 22532 MW: CD143BAD92B4C65 CRC64;

Query Match 72.5%; Score 642.5; DB 1; Length 203;
Best Local Similarity 65.0%; Pred. No. 1.1e-46;
Matches 134; Conservative 6; Mismatches 27; Indels 39; Gaps 3;

QY 4 MNPEVDYLFKLLIGDSGVKSCLLRFADDTYESYSTIGVDFKRTIELDGKTIKQ 63
    |||||
DB 1 MNPEVDYLFKLLIGDSGVKSCLLRFADDTYESYSTIGVDFKRTIELDGKTIKQ 60
QY 64 I-----ESFNNVKOMLOEIDRYASENNKILVG 91
    |||||
DB 61 IWDYAGQERFRITTSYRGAGHGIIVYDQESFNNVKOMLEIDRYASENNKILVG 120
QY 92 NKCDLTTRKKVVDYTTAKEPADSLGIPLETSAKNATNVEOSFMTAAEIKRMG---PG 147
    |||||
DB 121 NKSDLTJRKRVYSTDAKEFAESLGIPLETSAKNATNVEOAFMTAAEIKRMASCPVVP 180
QY 148 ATAGGAEKSNVAKIOSTPVKQSGGCC 173
    |||
DB 181 KPGQPVVRPT---EGKPINNKSSCC 203

RESULT 10
YPT1_PHTIN
ID YPT1_PHTIN STANDARD; PRT; 201 AA.
AC 001890;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Ras-like GTP-binding protein YPT1.
GN YPT1.
OS Phytothora infestans (potato late blight fungus).
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytothora.
OX NCBI_TaxID=4787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 52009;
RX MEDLINE=97128773; PubMed=8973313;
RA Chen Y., Roxby R.;
RT "Characterization of a Phytothora infestans gene involved in
RT vesicle transport.";
RL Gene 181:89-94(1996).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U30474; AAA40355.1; -.
DR HSSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_transfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PFO0071; ras; 1.
DR PRINTS: PR00449; RASTRNSPRMG.
DR SMART: SM00175; RAB; 1.

```

```

DR  TIGR00231: small_GTP; 1.
KW  GTP-binding; Lipoprotein; Prenylation; Protein transport;
KM  Multigene family.
FT  NP_BIND 15 22 GTP (BY SIMILARITY).
FT  NP_BIND 63 67 GTP (BY SIMILARITY).
FT  NP_BIND 121 124 GTP (BY SIMILARITY).
FT  DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT  LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
FT  LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
SQ  SEQUENCE 201 AA: 22292 MW: 89E553A693C8A65 CRC64;

Query Match 71.3%; Score 631.5; DB 1; Length 201;
Best Local Similarity 67.5%; Pred. No. 9.2e-46;
Matches 139; Conservative 8; Mismatches 18; Indels 41; Gaps 5;

QY 4 MNPEVDYLFKLLIGDSGVKSCLLRFADDTYESYSTIGVDFKRTIELDGKTIKQ 63
    |||||
DB 1 MNPEVDYLFKLLIGDSGVKSCLLRFADDTYESYSTIGVDFKRTIELDGKTIKQ 60
QY 64 I-----ESFNNVKOMLOEIDRYASENNKILVG 91
    |||||
DB 61 IWDYAGQERFRITTSYRGAGHGIIVYDQESFNNVKOMLEIDRYASENNKILVG 120
QY 92 NKCDLTTRKKVVDYTTAKEPADSLGIPLETSAKNATNVEOSFMTAAEIKRMGPGATAG 151
    |||||
DB 121 NKSDLTJRKRVYSTDAKEFAESLGIPLETSAKNATNVEKAFMMAAQIKRM---ANAP 177
QY 152 GAERSNVAKIOSTPVKQ--SGGQ--CC 173
    |||
DB 178 VAPKAGVRL--TPGQGVPSNGSKCC 201

RESULT 11
ARA5_ARATH
ID ARA5_ARATH STANDARD; PRT; 258 AA.
AC P28188;
DT 01-DEC-1992 (Rel. 24, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein ARA-5.
GN ARA-5 OR ATG02130 OR T7123.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotys; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altarfi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzilli A.,
RA Miltner J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Ventler J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE OF 64-258 FROM N.A.
RC STRAIN=Various strains; TISSUE=leaf;
RX MEDLINE=92084144; PubMed=1748311;
RA Anal T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,

```

RA Matsui M.;
 RT "Isolation and analysis of cDNAs encoding small GTP-binding proteins
 of Arabidopsis thaliana.";
 RL Gene 108:259-264(1991).
 CC -1- SIMILARITY: NO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U89959; AAC24370.1; -
 DR EMBL: D01027; BAA00832.1; -
 DR PIR: P50279; P50279.
 DR HSSP: P05713; 3RAB.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001806; Ras_trnsfmg.
 DR Pfam: PF00071; ras; 1.
 DR PRINTS: PR00449; RASTRNSFRNG.
 DR SMART: SM00175; RAB; 1.
 DR TIGRfams: TIGR00231; small_GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Multigene family.
 FT NP_BIND 70 77 GTP (BY SIMILARITY).
 FT NP_BIND 118 122 GTP (BY SIMILARITY).
 FT NP_BIND 176 179 GTP (BY SIMILARITY).
 FT DOMAIN 92 100 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 235 253 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 256 256 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 258 AA; 29169 MW; 34D278A0C504638F CRC64;
 Query Match 69.9%; Score 619.5; DB 1; Length 258;
 Best Local Similarity 61.6%; Pred. No. 1.2e+44;
 Matches 15; Conservative 15; Mismatches 30; Indels 33; Gaps 2;
 QY 3 MNPEYDYLKLLIGSGVSKSCLLRFADDTYESYSTIGVDFKRTITELDKITKL 62
 DB 55 TMNPEYDYLKLLIGSGVSKSCLLRFSDSDYESYSTIGVDFKRTITELDKITKL 114
 QY 63 QI-----ESFNKQWLOEIDRYASENVNKL 90
 DB 115 QIMDTAGQERFRTTSSYRGAGHIIIVYDTDESEFNKQWLEIDRYASDNVNL 174
 QY 91 GNKCDLTKKVVDTTAKERADSLGIPLETSAKNATNVQSFMTAAELKKRMPGATA 150
 DB 175 GNSDULFNKVVATETAKAFADDEIGIPMETSAKDATNVQAFMAASIKERNASOPAG 234
 QY 151 GAESKNVKIOSTPVKSGGCGC 173
 DB 235 NNARPTVQIRGQPVNOK-NGCC 256
 RESULT 12
 YPT2_MAIZE
 ID YPT2_MAIZE STANDARD: PRT: 203 AA.
 AC 005737;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GTP-binding protein YPTM2.
 GN YPTM2.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Coleoptile;
 RX MEDLINE=92115746; PubMed=1731354;

RA Palme K., Diefenthal T., Vingron M., Sander C., Schell J.;
 RT "Molecular cloning and structural analysis of genes from Zea mays
 (L.) coding for members of the ras-related ypt gene family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:787-791(1992).
 CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: ITS EXPRESSION IS WEAK IN STEMS, HIGHER IN
 CC ROOTS, LEAVES AND COLEOPTILES, BUT HIGHEST IN FLOWERS.
 CC -1- PPM: THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND
 CC BIOLOGICAL FUNCTION.
 CC -1- SIMILARITY: NO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X63278; CAA44919.1; -
 DR PIR: B38202; B38202.
 DR HSSP: P05713; 3RAB.
 DR MaizeDB: 78605; -
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001806; Ras_trnsfmg.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; ras; 1.
 DR PRINTS: PR00449; RASTRNSFRNG.
 DR SMART: SM00175; RAB; 1.
 DR TIGRfams: TIGR00231; small_GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.
 FT NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 121 124 GTP (BY SIMILARITY).
 FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
 FT LIPID 200 200 PALMITATE (BY SIMILARITY).
 FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 203 AA; 22475 MW; E241326E7ACD1B8A CRC64;
 Query Match 69.8%; Score 618.5; DB 1; Length 203;
 Best Local Similarity 63.4%; Pred. No. 1.1e+44;
 Matches 128; Conservative 11; Mismatches 30; Indels 33; Gaps 2;
 QY 4 MNPEYDYLKLLIGSGVSKSCLLRFADDTYESYSTIGVDFKRTITELDKITKL 63
 DB 1 MNPEYDYLKLLIGSGVSKSCLLRFADSDYSDYSTIGVDFKRTITELDKITKL 60
 QY 64 I-----ESFNKQWLOEIDRYASENVNKL 91
 DB 61 IMDTAGQERFRTTSSYRGAGHIIIVYDTDESEFNKQWLEIDRYASDNVNL 120
 QY 92 NKCDLTKKVVDTTAKERADSLGIPLETSAKNATNVQSFMTAAELKKRMPGATA 151
 DB 121 NKSDULFNKVVATETAKAFADDEIGIPMETSAKNATNVQAFMAASIKERNASOPAA 180
 QY 152 GAESKNVKIOSTPVKSGGCGC 173
 DB 181 NNARPTVQIRGQPVNOK-TSCC 201
 RESULT 13
 RIC1_ORYZA
 ID RIC1_ORYZA STANDARD: PRT: 202 AA.
 AC P40392;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ras-related protein RIC1.
 GN RIC1.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```
OC Eubacteriaceae; Oryzae; Oryza.
OX NCBI_TaxID=4530.
RN
RE SEQUENCE FROM N.A.
RC STRAIN=CV, Yamahouchi; TISSUE=Callus;
RD MEDLINE=94009718; PubMed=8405471;
RA Uchimiya H., Kidou S., Anai T., Umeda M., Aotsuka S., Tsuge T.,
RA Kato A.;
RT "Molecular structure of ras-related small GTP-binding protein genes
RT of rice plants and GTPase activities of gene products in Escherichia
RT coli.";
RL FEBS Lett. 332:282-286(1993).
CC
CC -1- FUNCTION: POSSESS GTPASE ACTIVITY.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: S66150; AAB28535.1; -.
DR PIR: S38740; S38740.
DR HSSP: P01112; 1PLK.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRPFAMS: TIGR00231; small_GTP; 1.
DR GTP-binding; lipoprotein; Prenylation; Palmitate; Protein transport.
KT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABE).
FT LIPID 200 200 PALMITATE (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 202 AA; 22561 MW; 5A1CB5F9325B8699 CRC64;
Query Match 69.28; Score 613.5; DB 1; Length 202;
Best Local Similarity 61.98; Pred. No. 2.9e-44;
Matches 125; Conservative 12; Mismatches 32; Indels 33; Gaps 2;
QY 4 MNPEDYLFKLLIDSGVSKCLLRFADDTYTSISTIGVDFKIRTIELDKRTIKLQ 63
DB 1 MNPEDYLFKLLIDSGVSKCLLRFADDSYLSYSTIGVDFKIRTVQDQKRTIKLQ 60
QY 64 I-----ESFNNVQWLOEIDRYASBNKLLVG 91
DB 61 IWDTRAGQERFRTTSSYRGAGHIVYDVTDQESFNNVQWLOEIDRYASBNKLLVG 120
QY 92 NKCDLTKKVVDTYAKKFAADSLGIPLETSAKNTNVEOSFMTMAAEIKRKGPGATAG 151
DB 121 NKCDLAKNRVSVSEAGKALADEIGIEPLETSKADTNVERAKFAFTMAGETIKRMA5QGRTN 180
QY 152 GAESNKKIQSTFVKQSGGCC 173
DB 181 ASKPAFVQMPROGVAQO--SSCC 201
RESULT 14
YPTL YEAST STANDARD: PRT: 206 AA.
AC P01123.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GTP-binding protein YPT1 (Protein YP2).
GN YPT1 OR YP2 OR YP1038C.
OS Saccharomyces cerevisiae (Baker's yeast).
```

```
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN
RE SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RD MEDLINE=84093563; PubMed=6318115;
RA Gallwitz D., Donath C., Sander C.;
RT "A yeast gene encoding a protein homologous to the human c-ha/bas
RT proto-oncogene product.";
RL Nature 306:704-707(1983).
CC
CC SEQUENCE FROM N.A.
CC STRAIN=S288C / AB972;
CC MEDLINE=95400292; PubMed=7670463;
CC Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
CC Sasamura S.-I., Sasamura M., Tsuchiya Y., Soeda E., Yokoyama K.,
CC Yamazaki M., Tashiro H., Eki T.;
CC "Analysis of the nucleotide sequence of chromosome VI from
CC Saccharomyces cerevisiae.";
CC Nat. Genet. 10:261-268(1995).
CC
CC [3]
CC PALMITOYLATION.
CC MEDLINE=88296440; PubMed=3042385;
CC Molenaar C.M.T., Prange R., Gallwitz D.;
CC "A carboxyl-terminal cysteine residue is required for palmitic acid
CC binding and biological activity of the ras-related yeast YPT1
CC protein.";
CC EMBO J. 7:971-976(1988).
CC
CC [4]
CC MUTAGENESIS, AND POSSIBLE FUNCTION.
CC MEDLINE=88223356; PubMed=3286011;
CC Schmitz H.D., Puzicha M., Gallwitz D.;
CC "Study of a temperature-sensitive mutant of the ras-related YPT1 gene
CC product in yeast suggests a role in the regulation of intracellular
CC calcium in yeast suggests a role in the regulation of intracellular
CC cell 53:635-647(1988).
CC
CC [5]
CC MUTAGENESIS OF EFFECTOR REGION.
CC MEDLINE=91184123; PubMed=2009858;
CC Becker J., Tan T.U., Trepte H.-H., Gallwitz D.;
CC "Mutational analysis of the putative effector domain of the
CC GTP-binding YPT1 protein in yeast suggests specific regulation by a
CC novel GAP activity.";
CC EMBO J. 10:785-792(1991).
CC
CC -1- FUNCTION: YPT1 IS ESSENTIAL FOR CELL VIABILITY. THE LOSS OF YPT1
CC RESULTS IN CYTOSKELETAL AND MITOTIC LESIONS. YPT1 IS REQUIRED FOR
CC TRANSIT OF PROTEINS FROM THE ER THROUGH GOLGI COMPARTMENT.
CC -1- PPM: THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND
CC BIOLOGICAL FUNCTION.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: X00209; CA25036.1; -.
DR EMBL: D50617; BAA09201.1; -.
DR PIR: A01374; TVBYO2.
DR HSSP: P05713; 3RAB.
DR SGD: S0001856; YPT1.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRPFAMS: TIGR00231; small_GTP; 1.
DR GTP-binding; lipoprotein; Protein transport; Golgi stack;
DR Endoplasmic reticulum; Prenylation; Palmitate.
KV
```


Query Match	Similarity	Score	766.5	DB 4	Length	201
Best Local Similarity	77.2%	Pred	4.1e-54			
Matches	156	Conservative	6	Mismatches	7	IndeIs
						Gaps
						2

QY 121 TSANKATNVEQSFMTMAEIKRMGPGATAGAEKSNKIOSTPVKQSGGCC 173
 DB 89 TSANKATNVEQSFMTMAEIKRMGPGATAGAEKSNKIOSTPVKQSGGCC 141

RESULT 4
 018332 PRELIMINARY: PRT: 205 AA.

AC 018332
 DT 01-JAN-1998 (TEMBLrel. 05. Created)
 DT 01-JAN-1998 (TEMBLrel. 05. Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21. Last annotation update)
 DE RAB1 protein
 GN RAB1 OR CG3320
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE=20196006; PubMed-10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Brill J.F., Agayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Bailew R.M., Basu A.V., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck H., Brostein P., Brotter P.,
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laspo P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Svirskas R., Tector A.C., Stapleton M., Strong R., Sun E.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.O., Wang X.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhou Q., Zhou L.,
 RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhang G., Zhou Q., Zhou L.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Zhu X., Smith H.O.,
 RA "The genome sequence of Drosophila melanogaster.";
 RT Science 287:2185-2195(2000).
 RL [2]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON R; TISSUE=HEAD;
 RA MEDLINE=97228579; PubMed-9074639;
 RA Satoh A.K., Tokumura F., Ozaki K.,
 RT "Rab proteins of Drosophila melanogaster: novel members of the Rab-
 protein family."
 RL FEBS Lett. 404:65-69(1997).
 RC 1 SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; AB03734; AAF5873.1; -.
 DR EMBL; D84312; BAA21705.1; -. .

DR HSSP: P05713; 3RAB.
 DR FlyBase: FBgn0016700; Rab1.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001806; Ras_trnstrng.
 DR InterPro: IPR002078; Sig54_interact.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; Ras_1.
 DR PRINTS: PR00449; BASTRNSFRNG.
 DR SMART: SM00175; RAB; 1.
 DR TIGRFS: TIGR00231; sma11_gtp; 1.
 DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 DR GTP-binding; Lipoprotein.
 KW GTP-binding 205 AA; 22762 MW; 9CE2691205FAE343 CRC64;
 SQ SEQUENCE

Query Match 77.1%; Score 683; DB 5; Length 205;
 Best Local Similarity 69.4%; Pred. No. 2, 2e-47;
 Matches 143; Conservative 12; Mismatches 17; Indels 34; Gaps 3;

QY 1 MSSMPEYDYLFKLLIGDSGVKSCLLRFADDTYESYSTTGVDKIRTEIDGKTI 60
 DB 1 MSSVPEYDYLFKLLIGDSGVKSCLLRFADDTYESYSTTGVDKIRTEIDGKTI 60
 QY 61 KIQI-----ESFNNVKOMIQEIDRYASENVKL 88
 DB 61 KIQIDVTAGQERFTTSSYRGAGLIYYDCTDSEFNVKOMIEIERACENVKL 120
 QY 89 LVGNKCDITTKKVVDTYTAKEPADSLGIPLETSAKATNVEQSFMTMAEIKRMGPGA 148
 DB 121 LVGNKSDLTTRKKVDTYTAFAEAOGLGIPLETSAKSATNVEQAFMTMAEIKRMVGPS 180
 QY 149 TAGAEKSNVKI-QSTPVKQSGGCC 173
 DB 181 SATD-NASKVKIDQGRPVENTKSGCC 205

RESULT 5
 016143
 AC 016143 PRELIMINARY: PRT: 202 AA.
 DT 01-JAN-1998 (TEMBLrel. 05. Created)
 DT 01-JAN-1998 (TEMBLrel. 05. Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21. Last annotation update)
 DE Small GTP-binding protein.
 GN RAB8.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Uno T., Nakajima A., Ueno M., Aizono Y.,
 RT "Isolation of cDNA encoding small GTP-binding protein from Bombyx
 mori.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC 1-1 SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; AF013572; AAB67169.1; -.
 DR HSSP: P05713; RAB8; 1.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001806; Ras_trnstrng.
 DR InterPro: IPR002078; Sig54_interact.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; Ras_1.
 DR PRINTS: PR00449; BASTRNSFRNG.
 DR SMART: SM00175; RAB; 1.
 DR TIGRFS: TIGR00231; sma11_gtp; 1.
 DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 DR GTP-binding; Lipoprotein.
 KW GTP-binding 202 AA; 22364 MW; 0338E895D63CDFE4 CRC64;
 SQ SEQUENCE

Query Match 75.7%; Score 671; DB 5; Length 202;
 Best Local Similarity 66.5%; Pred. No. 2e-46;
 Matches 135; Conservative 18; Mismatches 16; Indels 34; Gaps 3;

Oy	4	MNPEVDYLFKLLIGDSGVGKSCLLIRFADDDYTESYSTIGVDKPIRIETLDGKTIKQ	63
Db	1	MNPEVDYLFKLLIGDSGVGKSCLLIRFADDDYTESYSTIGVDKPIRIETLDGKTIKQ	60
Oy	64	I-----ESFNMYKQWLEIDRYASENVKLLVG	91
Db	61	IWDVGAQGRFRITTSYRYGAGHIIIVDCFTDQSFNSVWKQWLEIEDRACDNVKKLLVG	12
Oy	92	NKCDLTTKKYVDYTTAKFEADSLGIPELFTSKNKNNTNBOSSWYMAELKKMGCGATG	15
Db	121	NKCDLTTKKYVDYTFSTAKYAEQGLGIPLETSEKSNSTNWEQAEWMAETAKARVGPST-G	17
Oy	152	GAESKNVKI-OSTPYKQSGGCGCC	173
Db	180	AAPAGHVAKIDQGGPIDTGKSSCC	202
RESULT 6			
O9UA06			
ID	09UA06	PRELIMINARY;	PRT; 205 AA.
AC	O9UA06		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)	
De	Hypothetical 22.5 kDa protein.		
GN	C39F7.4.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;		
OC	Rhabdilitidae; Pelodierinae; Caenorhabditis.		
OX	NCBI_Taxid=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RX	MEDLINE=99069613; PubMed=9851916;		
RA	None;		
RT	"Genome sequence of the nematode C. elegans: a platform for		
RT	investigating biology. The C. elegans Sequencing Consortium.";		
RL	Science 282:2012-2018(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RA	Schneet P., Magg1 L., Dubbelde C.;		
RT	"The sequence of C. elegans cosmid C39F7.";		
RL	Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.		
RP	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BRISTOL N2;		
RA	Waterston R.;		
RT	"Direct Submission.";		
RL	Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.		
CC	-1 SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.		
DR	EMBL: AF010310; AAC69218.1; -		
DR	HSSP: P01112; IPII.		
DR	InterPro: IPR003579; GTPase_Rab.		
DR	InterPro: IPR001806; Ras_trnsfrmg.		
DR	InterPro: IPR002078; Sig54_interact.		
DR	InterPro: IPR005225; Small_GTP.		
DR	Pfam: PF00071; ras; 1.		
DR	PRINTS: PR00449; RASTRNSFRMG.		
DR	SMART: SM00175; RAB; 1.		
DR	TIGRFAMs: TIGR00231; small1_GTP; 1.		
DR	PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN.1.		
DR	GTP-binding; Hypothetical protein; Lipoprotein.		
KW	SEQUENCE 205 AA; 22545 MW; 5221920FP05842D3 CMC64;		
Query Match	74.5%;	Score 660;	DB 5; Length 205;
Best Local Similarity	66.2%;	Pred. NO. 1.5e-45;	
Matches 137;	Conservative 16;	Mismatches 18;	Indels 36; Gaps
Oy	1	MSNMPEDYLFKLLIGDSGVGKSCLLIRFADDDYTESYSTIGVDKPIRIETLDGKTI	60
Db	1	MSNMPEDYLFKLLIGDSGVGKSCLLIRFADDDYTESYSTIGVDKPIRIETLDGKTI	60

```

OY      61 KQI-----ESPNNYKQWLOEDIRYASEVVKL 88
      |||
Db      61 KQIDWDTGQGEFFRITTSYNGAGIIYVDTQETFTNNYKQWLOEDIRACERAVKVL 120
OY      89 LVGNKCDLTTKRVVYTTAKERFADSLGIPFELTSARKNATNVQSEPMTMAELTKRMGPGA 148
      |||||
Db      121 LVGNKCDLTTAKRAVEFQAQADYAGQGLPFELTSARKSSSTNVQAEFLTMSFKSRMGPGVQ 180
OY      149 TAGGAEKSNVAKIQ-STPYK-QSGGGCC 173
      ||||
Db      181 GAGGA--PGVRITGSPQVADKKKSGGCC 205

RESULT 7
O9HET3
ID      O9HET3      PRELIMINARY:      PRT:      201 AA.
AC      O9HET3;
DT      01-MAR-2001 (TREMBLrel, 16, Created)
DT      01-MAR-2001 (TREMBLrel, 16, Last sequence update)
DT      01-JUN-2002 (TREMBLrel, 21, Last annotation update)
DE      Secretion related GTPase (SrgB).
GN      SrgB.
OC      Aspergillus niger.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC      Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
CX      NCBI_TaxID=5061;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=N402;
RA      Punt P.J., Selbooth B., Weenink X.O., van Zeijl C.M., Lenders M.,
RA      Koretschey C., Ram A.F., Montijn R., Kubicek C.P.,
RA      Van den Hondel C.A.;
RT      Identification and characterisation of a family of secretion related
RT      small GTPase encoding genes from the filamentous fungus Aspergillus
RT      niger: a putative SEC4 homologue is not essential for growth.";
RL      Submitted (JUL-2000) to the EMBL/Genbank/DDJ databases.
CC      -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR      EMBL: AJ278659; CAC17833.1; -.
DR      HSSP: P01112; 1PLJ.
DR      InterPro: IPR003579; GTPase_Rab.
DR      InterPro: IPR001806; Ras_trnsfmg.
DR      InterPro: IPR02078; S1954_Interact.
DR      InterPro: IPR005225; Small_GTP.
DR      Pfam: PF00071; ras; 1.
DR      PRINTS: PR00449; RASPRNSFRMG.
DR      SMART: SM00175; RAB; 1.
DR      TIGRFAMS: TIGR00231; small_gtp; 1.
DR      PROSITE: PS00675; SIGMAS4_INTERACT_1; UNKNOWN_1.
KW      GTP-binding; Lipoprotein.
SQ      SEQUENCE 201 AA; 22303 MW; 8DC1A5B0BEEDF378 CRC64;

Query Match      74.0%; Score 655.5; DB 3: Length 201;
Best Local Similarity 67.2%; Pred. No. 3.4e-45;
Matches 137; Conservative 12; Mismatches 18; Indels 37; Gaps 3

OY      4 NNPEDYLFKLLIDSGVGSCLLRFADPDYTESYISTIGVDFKIRITELDGKTIKIQ 63
      |||||
Db      1 NNPEDYLFKLLIDSGVGSCLLRFADPDYTESYISTIGVDFKIRITELDGKTIKIQ 60
OY      64 I-----ESFNNYKQWLOEDIRYASEVNVKLVG 91
      |
Db      61 IWDTAGQGEFFRITTSYNGAGICVYVDTQETFTNNYKQWLOEDIRATGCVNKLIVG 120
OY      92 NNCGLTTKRVVYTTAKERFADSLGIPELTSARKNATNVQSEPMTMAELTKRMGPGATAG 151
      |||
Db      121 NNSDMDKRVVYTTAKERFADSLGIPELTSARKNATNVQAEFLTMSFKSRMGPGVQ 177
OY      152 GAESKNVAKIQSTPYKQSG--GGCC 173
      |
Db      178 VNNKPTVQVGGQGVQSGAGGCC 201

```

```

RESULT 8
Q9PB77 PRELIMINARY: PRT: 201 AA.
AC 09PB77:
AC 09PB77: 2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE YPA.
GN Aspergillus awamori.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RA Saloheimo M., Wang H., Valkonen M., Vasara T., Huuskonen A., Ward M.,
RA Penttila M.,
RA The secretory genes ypt1/ypt4 and nsf1/nsf4 from the filamentous
RT fungal trichoderma reesei and Aspergillus niger: evidence for global
RT transcriptional regulation of the secretory pathway in Trichoderma
RT reesei.
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC EMBL: AF244545; AAF63333.1; -.
CC HSSP: P01112; 1ptl.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR002078; Sig54_Interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMNG.
DR SMART: SM00175; RAB; 1.
DR TIGRFS: TIGR00231; small_gtp; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KM GTP-binding; Lipoprotein.
SQ SNOUCE 201 AA; 22335 MW; 9D710F1ABE279 CRC64;

Query Match 73.8%; Score 653.5; DB 3; Length 201;
Best Local Similarity 67.2%; Pred. No. 56-45;
Matches 137; Conservative 12; Mismatches 18; Indels 37; Gaps 3;

AC 040203 PRELIMINARY: PRT: 202 AA.
AC 040203:
AC 040203: 1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE RAB1C.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Loteeae; Lotus.
OX NCBI_TaxID=34305;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE-ROOT NODES;
RX MEDLINE=97231679; PubMed=9076991;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RA "Identification of new protein species among 33 different small GTP-
RT binding proteins encoded by cDNAs from Lotus japonicus, and expression
RT of corresponding mRNAs in developing root nodules."
RT Plant J. 11:237-250(1997).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
RN [1]
RP SEQUENCE FROM N.A.
RA Saloheimo M., Wang H., Valkonen M., Vasara T., Huuskonen A., Ward M.,
RA Penttila M.,
RA The secretory genes ypt1/ypt4 and nsf1/nsf4 from the filamentous
RT fungal trichoderma reesei and Aspergillus niger: evidence for global
RT transcriptional regulation of the secretory pathway in Trichoderma
RT reesei.
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC EMBL: AF244545; AAF63333.1; -.
CC HSSP: P01112; 1ptl.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR002078; Sig54_Interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMNG.
DR SMART: SM00175; RAB; 1.
DR TIGRFS: TIGR00231; small_gtp; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KM GTP-binding; Lipoprotein.
SQ SNOUCE 202 AA; 22633 MW; 91633EBFAD07662F CRC64;

Query Match 72.7%; Score 644; DB 10; Length 202;
Best Local Similarity 65.8%; Pred. No. 2,9e-44;
Matches 133; Conservative 8; Mismatches 27; Indels 34; Gaps 3;

AC 09SEH3 PRELIMINARY: PRT: 202 AA.
AC 09SEH3:
AC 09SEH3: 2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE RAS-related small GTP-binding protein (RAS-related small GTP-binding
DE protein RAB1C) (At4g17350/DLA600C).
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIN; TISSUE=EPIDERMAL;
RA Bischoff F., Gode K., Palme J.;
RA "Arabidopsis involved in ER-Golgi traffic in meristematic cells in
RT Arabidopsis."
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk K., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,

```

RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamita A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.T., Sakurai T.,
RA Saito M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AF127134; AAF22133.1; -
DR EMBL: A161546; CAB78756.1; -
DR EMBL: AY052204; AAK97673.1; -
DR HSSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmrng.
DR InterPro: IPR002078; Sig54_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_gtp; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
KW VARIANT 70
SQ SEQUENCE 202 AA; 22318 MW; 954FA24C3110FC12 CRC64;

Query Match 72.5%; Score 642; DB 10; Length 202;
Best Local Similarity 65.8%; Pred. No. 4.2e-44;
Matches 133; Conservative 11; Mismatches 24; Indels 34; Gaps 3;

QY 4 MNPEYDYLFKLLIGDSGVGSKCLLRPADTYESTYSTIGVDFKRTIELDKTKLQ 63
DB 1 MNPEYDYLFKLLIGDSGVGSKCLLRPADSYLDSYSTIGVDFKRTIELDKTKLQ 60
QY 64 I-----ESFNNKQWLOEIDRYASENVKLLVG 91
DB 61 IWDTAGOERFRTITSSYRGAGHGIIVYDVTDLSEFNNKQWLOEIDRYASENVKLLVG 120
QY 92 NKCDLTTKKVVDYTTAKFEADSLGIPLETSAKNATNVEQSFMTMAEIKRKGATAG 151
DB 121 NKCDLTSKVVSTETAKFADSLGIPLETSAKNATNVEAFMTAIAIKTRMA-SOPAG 179
QY 152 GAESKNVAKIOSTPVKSGGCC 173
DB 180 GSKPPTVQIRGQPVNQ-SGCC 200

RESULT 11
Q9HDT5 PRELIMINARY; PRT; 202 AA.
AC Q9HDT5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Small GTP-binding protein YPT1.
GN YPT1.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
CC NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RTCC-30;
RA Salohelmo M.L.A., Wang H., Valkonen M., Vasara T., Huuskonen A.,
RA Rikonen M., Pakula T., Ward M., Penttilae M.;
RT "The secretory genes ypt1/ypt4 and nsf1/nsf4 from the filamentous
fungi Trichoderma reesei and Aspergillus niger var. awamori: evidence
RT for global transcriptional regulation of the secretory pathway in T.
reesei";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: AJ277108; CAC17744.1; -

DR HSSP: P01112; 1PLJ.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmrng.
DR InterPro: IPR002078; Sig54_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_gtp; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 202 AA; 22421 MW; A60E10E2AB1B2E7 CRC64;

Query Match 72.2%; Score 640; DB 3; Length 202;
Best Local Similarity 65.7%; Pred. No. 6.1e-44;
Matches 134; Conservative 13; Mismatches 21; Indels 36; Gaps 4;

QY 4 MNPEYDYLFKLLIGDSGVGSKCLLRPADTYESTYSTIGVDFKRTIELDKTKLQ 63
DB 1 MNPEYDYLFKLLIGDSGVGSKCLLRPADTYESTYSTIGVDFKRTIELDKTKLQ 60
QY 64 I-----ESFNNKQWLOEIDRYASENVKLLVG 91
DB 61 IWDTAGOERFRTITSSYRGAGHGIIVYDVTDMSEFNNKQWLOEIDRYATEGVNKLVG 120
QY 92 NKCDLTTKKVVDYTTAKFEADSLGIPLETSAKNATNVEQSFMTMAEIKRKGATAG 151
DB 121 NKSDMSDKVVEYTTAKFEADSLGIPLETSAKNATNVEQAFILMAQIKERM-GTTTA 178
QY 152 GAESKNVAKI-OSTPV-KOSGGGCC 173
DB 179 NNTKPSVHVGGGQGVGNSSSSCC 202

RESULT 12
Q08154 PRELIMINARY; PRT; 203 AA.
AC Q08154;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE GTP-binding protein.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaceae; Pisum.
CC NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RX MEDLINE=94290824; PubMed=8019783;
RA Nagano Y., Mural N., Matsuno R., Sasaki Y.;
RT "Isolation and characterization of cDNAs that encode eleven small GTP-
binding protein from Pisum sativum";
RL Plant Cell Physiol. 34:447-455(1993).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: D12549; BAA02117.1; -
DR HSSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfmrng.
DR InterPro: IPR002078; Sig54_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; Ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_gtp; 1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 203 AA; 22463 MW; 6F2E1007A31EF34 CRC64;

Query Match 71.8%; Score 636.5; DB 10; Length 203;
Best Local Similarity 64.9%; Pred. No. 1.2e-43;
Matches 131; Conservative 11; Mismatches 27; Indels 33; Gaps 2;

RESULT 13
DB 4 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 63
DB 1 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 60
DB 64 I-----ESFNVMKQMLEIDRYASBNVKNLLVG 91
DB 61 IMDTAGOEFRFTTSSYRGANGIIVYDVTDESEFNVMKQMLEIDRYASBNVKNLLVG 120
DB 92 NKCDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 151
DB 121 NKDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 180
DB 152 GAESNVKIQSTPVKQSGGCC 173
DB 181 GAKPPTVQIRGQPVNQSG -GCCC 201

RESULT 14
DB 4 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 63
DB 1 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 60
DB 64 I-----ESFNVMKQMLEIDRYASBNVKNLLVG 91
DB 61 IMDTAGOEFRFTTSSYRGANGIIVYDVTDESEFNVMKQMLEIDRYASBNVKNLLVG 120
DB 92 NKCDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 151
DB 121 NKDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 180
DB 152 GAESNVKIQSTPVKQSGGCC 173
DB 180 GAKPPTVQIRGQPVNQSG -GCCC 200

RESULT 14
DB 4 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 63
DB 1 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 60
DB 64 I-----ESFNVMKQMLEIDRYASBNVKNLLVG 91
DB 61 IMDTAGOEFRFTTSSYRGANGIIVYDVTDESEFNVMKQMLEIDRYASBNVKNLLVG 120
DB 92 NKCDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 151
DB 121 NKDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 180
DB 152 GAESNVKIQSTPVKQSGGCC 173
DB 180 GAKPPTVQIRGQPVNQSG -GCCC 200

RESULT 14
DB 4 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 63
DB 1 MNPEDYLFKLLIGDSGVKSCLLRFADDTYESTYITGVDFKRTIELDGTITKLO 60
DB 64 I-----ESFNVMKQMLEIDRYASBNVKNLLVG 91
DB 61 IMDTAGOEFRFTTSSYRGANGIIVYDVTDESEFNVMKQMLEIDRYASBNVKNLLVG 120
DB 92 NKCDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 151
DB 121 NKDDLTTRKVVYDTTAKAFADSLGIPLETSAKNATNVEOSFPMMAEIKRMRGPGATAG 180
DB 152 GAESNVKIQSTPVKQSGGCC 173
DB 180 GAKPPTVQIRGQPVNQSG -GCCC 200

```

RESULT 15
024112 ID 024112 PRELIMINARY; PRT; 203 AA.
AC 024112;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Small GTP-binding protein.
OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID:4092;
RN [1]
RP SEQUENCE FROM N.A.
RA Borisjuk N., Stalio L., Kaydamov C., Senger S., Tewes A.,
RA Mantoufel R.;
RT "Cloning and characterization of cDNA clones differentially expressed
RT during somatic embryogenesis of Nicotiana plumbaginifolia: a mRNA
RT differential display approach."
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL: Y08425; CAA69701.1; -.
DR HSP; P01112; 1PL.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR001806; Ras_trnsfmg.
DR InterPro: IPR002078; Sig54_interact.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM0175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_gtp; 1.
DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 203 AA; 22516 MW; 13799234FF719DE7 CRC64;

Query Match 71.4%; Score 632.5; DB 10; Length 203;
Best Local Similarity 63.9%; Pred. No. 2.5e-43;
Matches 129; Conservative 12; Mismatches 28; Indels 33; Gaps 2;

QY 4 MNPEDYLFKLLIGDSGVKSCLLRPADPTYESYSTIGVDPKIRTIELDGKTIKQ 63
DB 1 MNPEDYLFKLLIGDSGVKSCLLRPADPTYESYSTIGVDPKIRTVQDGTIKIQ 60
QY 64 I-----ESFNVKQWLQEIIDRYASENVKILYG 91
DB 61 IWDTAGGERFRRTTSSYIRGAHGIIVYDVTDESFNVNKQWLSEIDRYASDSVKNILYG 120
QY 92 NKCDLITTKVVDYTTAKFADSLGIPLETSAKNATNVEQSFTMTAAEIKKRMGPGATAG 151
DB 121 NKCDLITKQKVVSTETAGAFADSLGIPLETSAKNATNVEQAFMAASIKKRMASQPSAN 180
QY 152 GAESKNVYKIOSTPVKQSGGGCC 173
DB 181 NARPPTVQIRGOVPYNOK-SGCC 201

```

Search completed: November 17, 2002, 13:06:53
Job time : 83 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame-plus-p2n model

Run on: November 17, 2002, 13:06:58; Search time 1959 Seconds

(without alignments)
1430.229 Million cell updates/sec

Title: US-09-820-003a-2
Perfect score: 886
Sequence: 1 MSHNMPEDYLFKLLIGDS.....EKSNNKIQSPVKNQSGGCC 173

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+P2n.model -DEV-xih
-O/cgn2.1/USPTO.spool/US09820003/runat.1112002.134616.3189/app-query.fasta.1.327
-DB-EST -OPMT-fastap -SUFFIX-est -MINMATCH=0.1 -LOOPEXT=0
-UNIT-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORP=500 -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE-LOCAL
-OUTFMT=ptc -NOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09820003.cgn.1.1.899.8runat.1112002.134616.3189 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGOQUERY -NEG.SCORES=0 -NATF -LONGLOG -DEV.TIMEOUT=120
-NARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em.estba:*
2: em.esthum:*
3: em.estlin:*
4: em.estm:*
5: em.estoy:*
6: em.estol:*
7: em.estro:*
8: em.hic:*
9: gb.est1:*
10: gb.est2:*
11: gb.hic1:*
12: gb.est3:*
13: gb.est4:*
14: gb.est5:*
15: em.estlum:*
16: em.estom:*
17: gb.gss:*
18: em.gss.hum:*
19: em.gss.liv:*
20: em.gss.pla:*
21: em.gss.vit:*
22: em.gss.fun:*
23: em.gss.man:*
24: em.gss.mus:*
25: em.gss.other:*
26: em.gss.pri:*
27: em.gss.tod:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860	97.1	948	9 AL539022	AL539022 AL539022
2	860	97.1	948	14 BM919860	BM919860 AGENCOURT
3	849	95.8	1130	13 BM542833	BM542833 AGENCOURT
4	847	95.6	862	14 BM223273	BM223273 AGENCOURT
5	847	95.6	944	14 BM223273	BM223273 AGENCOURT
6	831.5	93.8	646	13 BM2093078	BM2093078 AGENCOURT
7	825	93.1	957	14 BM219698	BM219698 AGENCOURT
8	822	92.8	864	14 BM437146	BM437146 AGENCOURT
9	820	92.6	901	14 BM452262	BM452262 AGENCOURT
10	819	92.4	1105	13 BM452262	BM452262 AGENCOURT
11	814	91.9	690	10 BM954378	BM954378 EST366448
12	809	91.3	945	9 AL530265	AL530265 AL530265
13	808.5	91.3	642	13 BM7073109	BM7073109 AGENCOURT
14	807	91.1	872	13 BM684554	BM684554 AGENCOURT
15	806	91.0	855	13 BM739625	BM739625 AGENCOURT
16	801	90.4	807	12 BM698150	BM698150 AGENCOURT
17	800	90.3	696	14 BM786746	BM786746 K-EST0065
18	800	90.3	890	14 BM437514	BM437514 AGENCOURT
19	795	89.7	673	12 BM692590	BM692590 AGENCOURT
20	795	89.7	846	12 BM619962	BM619962 AGENCOURT
21	794	89.6	1067	12 BM922791	BM922791 AGENCOURT
22	794	89.6	909	12 BM743352	BM743352 AGENCOURT
23	788.5	89.0	859	13 BM1599480	BM1599480 AGENCOURT
24	788.5	89.0	1034	14 BM067968	BM067968 AGENCOURT
25	788	88.9	611	9 AL048520	AL048520 un61106.1
26	787	88.8	828	13 BM652154	BM652154 AGENCOURT
27	783	88.4	839	13 BM858729	BM858729 AGENCOURT
28	783	88.4	913	12 BM970841	BM970841 AGENCOURT
29	780	88.0	933	13 BM681067	BM681067 AGENCOURT
30	778.5	87.9	588	13 BM74071	BM74071 AGENCOURT
31	777	87.7	901	13 BM450994	BM450994 AGENCOURT
32	777	87.7	937	13 BM138727	BM138727 AGENCOURT
33	776	87.6	676	10 BM913399	BM913399 AGENCOURT
34	775.5	87.5	612	10 BM959537	BM959537 AGENCOURT
35	775	87.5	1077	13 BM476507	BM476507 AGENCOURT
36	769.5	86.9	881	14 BM424046	BM424046 AGENCOURT
37	769	86.8	1047	14 BM800733	BM800733 AGENCOURT
38	768.5	86.7	643	9 AL535937	AL535937 AGENCOURT
39	768.5	86.7	744	13 BM045935	BM045935 AGENCOURT
40	768.5	86.7	766	12 BM826700	BM826700 AGENCOURT
41	768.5	86.7	821	9 BM75331	BM75331 AGENCOURT
42	768.5	86.7	825	9 AL542311	AL542311 AGENCOURT
43	768.5	86.7	826	10 BM561888	BM561888 AGENCOURT
44	768.5	86.7	837	9 AL520853	AL520853 AGENCOURT
45	768.5	86.7	862	14 BM941767	BM941767 AGENCOURT

ALIGNMENTS

RESULT 1
AL539022
LOCUS AL539022 LTR.FL013.FBrrn1 948 bp mRNA linear EST 16-FEB-2001
DEFINITION AL539022 LTR.FL013.FBrrn1 Homo sapiens cDNA clone CSDF030YG06 5
ACCESSION AL539022
VERSION AL539022
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 948
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="F013.FBRL1"
/dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"
/lab_host="DH10B"
/note="Organ: Fetal brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com url : http://fulllength.invitrogen.com"
BASE COUNT 277 a 187 c 239 g 244 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 1,87e-94 Length: 948
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best Local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
Gaps: 1

US-09-820-003a-2 (1-173) x AL539022 (1-948)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
Db 166 ATGTCAGCATGATCCGCAATATGATTATTATCAAGTACTCTGATTGGCGACTCA 225
QY 21 GlyValGlySerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40
Db 226 GGGGTGGAAAGCTTGGCTTCTTCTTGAAGTTCACATATACATATACAGAAAGCTAC 285
QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
Db 286 ATCAGCACATTTGGTGTGGATTCCAAATAGAACTATAGATTAGACGGGAAACATC 345
QY 61 LysLeuGlnIle----- 64
Db 346 AAGCTCAATATATGGACACAGCAGCCAGAAAGATTTCGAACATCACCTCCAGTTAT 405
QY 65 -----GluSerPheAsn 68
Db 406 TACAGAGAGCCCATGCGATCATAGTTGTATGATGTGACAGATCAGAGGTCTTCAAT 465
QY 69 AsnValLysGlnTrpLeuGlnGluIleAspArgTyrIleAspSerGluAsnValAsnLysLeu 88
Db 466 AATGTTAAACAGTGGCTGCGCAAGAAATAGATCGTTATGCCGTAAGAAATGTCAACAAATTG 525
QY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108
Db 526 TTGGTAGAGAACAAATGTATGTGACCAACAAGAAAGTACTAGACTACACAGCGCAAG 585
QY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
Db 586 GAATTTGCTATTCCCTTGGAAATTCGTTTGGAAACAGTCTTAAGATGCAACCAAT 645
QY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
Db 646 GTGAAACAGCTTTTCATGACGATGCGCAGCTGAGATTAAAAAGCAATGGGTCCCGGAGCA 705

QY 149 ThrAlaGlyValAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
Db 706 ACAGCTGGTGGTGGTGGAGAGTCCCAATGTTAAATTCAGAGCAGCTCAGTCAAGCAGTCA 765
QY 169 GlyValGlyLysCys 173
Db 766 GGTGGAGGTTCGTCG 780

RESULT 2
BM919860 976 bp mRNA linear EST 12-MAR-2002
LOCUS BM919860
DEFINITION AGENCOUNT_6708189 NIH_MGC_120 Homo sapiens CDNA IMAGE:5749635
5', mRNA sequence.
ACCESSION BM919860
VERSION BM919860.1 GI:19370239
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 976)
AUTHORS NIH-MGC http://img.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12779 row: f column: 04
High quality sequence stop: 757.
Location/Qualifiers

FEATURES
source
1. 976
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: Pooled pancreas and spleen; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (Rcovr) site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT 285 a 196 c 246 g 248 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.95e-94 Length: 976
Score: 860.00 Matches: 173
Percent Similarity: 84.39% Conservative: 0
Best Local Similarity: 84.39% Mismatches: 0
Query Match: 97.07% Indels: 32
Gaps: 1

US-09-820-003a-2 (1-173) x BM919860 (1-976)

QY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
Db 177 ATGTCAGCATGATCCGCAATATGATTATTATCAAGTACTCTGATTGGCGACTCA 226
QY 21 GlyValGlySerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40
Db 237 GGGGTGGAAAGCTTGGCTTCTTCTTGAAGTTCACATATACATATACAGAAAGCTAC 296
QY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13344 row: k column: 10
 High quality sequence stop: 728.
 Location/Qualifiers

FEATURES

source

1..862
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6065561"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Testis; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."

BASE COUNT 254 a 170 c 228 g 210 t
 ORIGIN

Alignment Scores:

Pred.: No.: 6.26e-93 length: 862
 Score: 847.00 Matches: 173
 Percent Similarity: 83.98% Conservative: 0
 Best Local Similarity: 83.98% Mismatches: 0
 Query Match: 95.60% Indels: 33
 Gaps: 14

US-09-820-003a-2 (1-173) x BQ223273 (1-862)

QY 1 MetSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
 Db 179 ATGCGCAGCATGATCCCGAATATGATTATTATTCAGTACTTGTGATGGCGACTCA 238
 QY 21 GlyValGlyLysSerCysLeuLeuLeuArgPheAlaAspThrTyrThrGluSerTyr 40
 Db 239 GGGGTGGGAAGCTTCCCTCTTCTTGGTGCACATGATATACAGAAAGCTAC 298
 QY 41 ILSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
 Db 299 ATACGCAATTTGGTGTGATTTCAAAATAGAACTATAGAGTAGAGGGAACAATC 358
 QY 61 LysLeuGlnIle----- 64
 Db 359 AAGCTTAAATATATGGACACAGCAGCCAGGAAAGATTTCGAACATCACCCTCAATTAT 418
 QY 65 -----GluSerPheAsn 68
 Db 419 TACAGAGACCCATGCGCATCATAGTTGTATGATGATGACAGATTCAGAGTCCCTTCAT 478
 QY 69 AspValLysGlnIlePheLysGlnIleAspArgTyrIleAspGluAsnValAspLysLeu 88
 Db 479 AATGTTAAAGTGGTGGCAGGAATATGATGCTTATCCAGTGAATGTCACAAATTTG 538
 QY 89 LeuValGlyLysLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108
 Db 539 TTGGTAGGAACAATGTGATCTGCACCAAGAAGTAAGTACAGCTACACACAGCGAAG 598
 QY 109 GluPheAlaAspSerLeuGlyIleProPheLysGlnIleThrSerAlaLysAsnIleAsn 128
 Db 599 GAATTTGCTATTCCTCTGGAATTCCTTTTGGAAACAGTGTAGATGCAACGAAT 658
 QY 129 ValGluGlnSerPheMetThrMetAlaAlaGlnIleLysLysArgMetGlyProGlyAla 148
 Db 659 GTAGAACAGTCTTTCATGAGATGCGAGCTGAGATTAAAGGCAATGGGTCCCGAGACA 718

QY 149 ThrAlaGlyLysAlaGlnLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
 Db 719 ACAGCTGGTGGTGGTGGAGAGTCCCAATGTAAATTCAGACACTCCAGTCAAGCAGCTCA 778
 QY 169 GlyGlyGlyCysCys 173
 Db 779 GGTGGAGGTGGCTGC 794

RESULT 5

BQ926234

LOCUS BQ926234

DEFINITION AGENCOURT_8801639 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6315467

ACCESSION BQ926234

VERSION BQ926234.1 GI:22341265

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Crustacea; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 944)

JOURNAL NIH-MGC <http://mgs.nci.nih.gov/>

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Susan L. Sullivan, PhD.

cDNA Library Preparation: Resgen, Invitrogen Corp

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM13745 row: n column: 12

High quality sequence stop: 750.

FEATURES

source

1..944
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:6315467"
 /clone_lib="NIH_MGC_129"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Olfactory epithelium; Vector:
 pCMV-SPORT6.1.cdb; Site_1: EcoRV; Site_2: NotI; Cloned
 unidirectionally. Primer: Oligo dT. Average insert size
 2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this
 is a NIH-MGC Library."

BASE COUNT 282 a 182 c 211 g 266 t 3 others
 ORIGIN

Alignment Scores:

Pred. No.: 7.2e-93 length: 944
 Score: 847.00 Matches: 170
 Percent Similarity: 84.16% Conservative: 0
 Best Local Similarity: 84.16% Mismatches: 0
 Query Match: 95.60% Indels: 32
 Gaps: 14

US-09-820-003a-2 (1-173) x BQ926234 (1-944)

QY 4 MetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSerGlyValGly 23
 Db 20 ATGAATCCGGAATATGATTATTATTCAGTTACTTCTGATTTGGCGATTCTGGGGTTGGA 79
 QY 24 LysSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyrIleSerThr 43
 Db 80 AAGTCTGCTCTTCTCTTGAAGTTGAGATGATACGATACGGAAGCTTACATCAGCACA 139
 QY 44 ILleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIleLysLeuGln 63
 Db 140 ATTGGTGGATTTCAGATACGACTATAGAGTTAGATGGAGAAACATCAAGCTACAG 199
 QY 64 Ile----- 64

```

Db 200 ATATGGACACAGACAGCCAGGAAAGATTTCAGACATTCCTCATTTATACAGAGA 259
|||
QY 65 -----GUSERPHEASNAVALYS 71
Db 260 GCCATGCAATCAATAGTGTATGATGATGACAGATCAGAGCTCTTCAATAAAGCTTAAA 319
QY 72 GATTPLEUGINGLUILEASPARGTYRALASERGIUAENVALISANLYSLEUENVALIGLY 91
Db 320 CAGTGGCTGACAGAGATGATGCTTACGCTACGCTGAAATGTCAAAAGTTGGTGGG 379
QY 92 ASNLVSCYASPLEUTHRTHRLYSVALVALASPTYRTHRLTHRALYSGLUPHEALA 111
Db 380 AACCAATGATACCTGACCAACAAAGAGTAGAGCTACAAACAGCAAGAGATTTGCA 439
QY 112 ASPSERLEUGIYLEPROPHLEUGIUTHRSERIALALYSASNAIATHRSNAVALIGL 131
Db 440 GATTCCTTGAAATCCATTTTGGAAACAGCTGCTAAGAACCAAGATGTAGAACAG 499
QY 132 SERPHEKETHMETALALAGLUILELYSLYSATGMETGLYPROGLYALATHRALIGLY 151
Db 500 TCTTTCATGACGATGACGATGAGATTAAGCAATGGGCTCTGAGCTTCAGCTGGT 559
QY 152 GLYALAGIULYSERASNAVALYSLIEGLINSERTHRPROVALYGLINSERGLYGLY 171
Db 560 GGTCCGAGAAATCCATGTTAAATCCAGAGCAGCTCAGTCAAGCAGTCAAGGAGGAG 619
QY 172 CYSCYS 173
Db 620 TGCTGC 625

RESULT 6
LOCUS B0219698 957 bp mRNA linear EST 02-MAY-2002
DEFINITION B0219698 NIBB Mochii normalized Xenopus early gastrula library
Xenopus laevis cDNA clone Xl141f18 5', mRNA sequence.
ACCESSION B0219698
VERSION B0219698.1 GI:17592487
KEYWORDS EST
SOURCE B0219698.1 GI:17592487
ORGANISM African clawed frog.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodidae; Xenopus.
REFERENCE 1 (bases 1 to 646)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara
,T.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasi Shin-I
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
FEATURES
source
1. 646
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_id="Xl141f18"
/clone_lib="NIBB Mochii normalized Xenopus early gastrula
library"
/tissue_type="whole embryo"
/dev_stage="stage 10.5"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Newkooop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT 208 a 129 c 161 g 148 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3,16e-91 Length: 646
Score: 831.50 Matches: 169
Percent Similarity: 82.93% Conservative: 1
Best Local Similarity: 82.44% Mismatches: 2
Query Match: 93.85% Indels: 33
DB: 13 Gaps: 2

US-09-820-003a-2 (1-173) x B0219698 (1-646)
QY 1 MetSerSerMetAsnProGluuTyRAspTyRLeuPheLysLeuLeuLeuIleGlyAspSer 20
Db 35 ATGTCCACATGATGATCCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 94
QY 21 GLYVALIGLYYSERCYSLLEUENVALARGPHEALASPARTRTYRTHRLGISEUETFR 40
Db 95 GGTGGGGAAGACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 154
QY 41 ILESERTHRIIEGLYVALASPHLYSLIARATHRILEGLULLEUASPGIYSTRILLE 60
Db 155 ATCAGTACCAATGTGTGTTGATTTTAAATCAGACTATGAGTTAGATGAGTAAAAACATTC 214
QY 61 LYSLEUGLILE----- 64
Db 215 AAACGTGCAATTTTGGAGACATGCGCGCCAGAGAAAGATTTCGAACCATTCATCAAGTTAC 274
QY 65 -----GUSERPHEASN 68
Db 275 TACAGGGAGCCGACGCGCATCTGTTGTGTATGACGTACAGACGAGATCTCTTCAAC 334
QY 69 ASNAVALYSGLINTPLEUGINGLUILEASPARGTYRALASERGIUAENVALISANLYSLEU 88
Db 335 AATGTCACAGATGCTGCTTACGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 394
QY 89 LEUVALIGLYASNYSCYASPLEUTHRTHRLYSVALVALASPTYRTHRLTHRALYS 108
Db 395 TTGGTAGCACAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 454
QY 109 GLUPHEALASPSERLEUGIYLEPROPHLEUGIUTHRSERIALALYSASNAIATHRSN 128
Db 455 GAGTTTGACAGACTCTCTGGGAATTCATTTTGGAAACAGCGCAAGAACCGGCAACAC 514
QY 129 VALIGLINSERPHMETHRMETALALAGLUILELYSLYSATGMETGLYPROGLYALA 148
Db 515 GTTAGGACAGGCGCTTCATGACAAATGACAGCGAGATCAAAAAGCAATGGGCGCTGCGCC 574
QY 149 THRALAGIYGLIAGLUILYSERASNAVALYSLIEGLINSERTHRPROVALYSLINSER 168
Db 575 ACAGAGGCGGCTCAAGAGAG--AACGTCAAAATCCAGAGCAGCTCAGTCAAGAGCTCT 631
QY 169 GLYGLYGLYCYSCYS 173
Db 632 AGTGGAGGATGCTGC 646

RESULT 7
LOCUS B0219698 957 bp mRNA linear EST 02-MAY-2002
DEFINITION B0219698 NIBB Mochii normalized Xenopus early gastrula library
Xenopus laevis cDNA clone Xl141f18 5', mRNA sequence.
ACCESSION B0219698
VERSION B0219698.1 GI:20401098
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 957)
AUTHORS NIH-MGC http://mgi.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rtm.nhl.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.

```

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM13305 row: e column: 14
High quality sequence stop: 671.

FEATURES
source
location/Qualifiers
1..957
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6051445"
/clone_lib="NCI_CGAP-St1"
/lab_host="DH10B (71-resistant)"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 269 a 208 c 250 g 229 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	3,59e-90	Length:	957
Score:	825.00	Matches:	171
Percent Similarity:	83.41%	Conservative:	0
Best Local Similarity:	83.41%	Mismatches:	2
Query Match:	93.12%	Indels:	33
DB:	14	Gaps:	1

US-09-820-003a-2 (1-173) x BQ219698 (1-957)

```
OY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
DB 200 ATGTCCAGCATGATGCCGGAATATGATTATTATTCAGTTACTTCTGATTGGGATCT 259
OY 21 GlyValGlySerSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40
DB 260 GGGGTGGGAAGTCTCCCTTCCTTAGGTTGCAGATGATACGTATACGGAAGCTAC 319
OY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 320 ATACGACCAATTGGTGTGATTTCACATACGACATACGATACGATGAGTGAAGAAACATC 379
OY 61 LysLeuGlnIle----- 64
DB 380 AAGCTACAGATATGGACACAGCAGGCCAGAAAGATTGCAACAATCCACTCCAGTTAT 439
OY 65 -----GluserPheAsn 68
DB 440 TACAGAGAGCCCATGGCATCATAGTTGTGTATGATGTACAGATCAGAGTCCCTTCAT 499
OY 69 AsnValIleGlnTrpLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
DB 500 AACGTTAAACAGTGGCTGCGAGAGATGATGCTACGCCGAGTGAATATGCAACAAGTTG 559
OY 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108
DB 560 TTGGTAGGGAACAATGTGACTGACCAACAAGAAAGTACTAGACTACACAGCAGCAAG 619
OY 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
DB 620 GAATTGGCAATTCCTCTGGAAATTCATTTTGGAAACCGTCTAAGAACGCAACCAAT 679
OY 129 ValGluGlnSerPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
DB 680 GTTGAACAGCTCTTTCATGACGATGGCAGCTGACATTANAAAGCAAGGGTCTCGGAGCT 739
OY 149 ThrAlaGlyGlyAlaGluLysSerAsnValLysIleGlnSerThrProValLysGlnSer 168
DB 740 ACAGCTGTGTGTGCGGAGAAATCCAT-GTTAAATCCAGAGCACTCCAGTCAAGCAATCA 798
OY 169 GlyGlyGlyCysCys 173
```

Db 799 GGTEGAGGCTGCTGC 813

RESULT 8 BQ437146 864 bp mRNA linear EST 24-MAY-2002
LOCUS BQ437146
DEFINITION AGENCOURT_7909781 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6156281
5', mRNA sequence.

ACCESSION BQ437146
VERSION BQ437146.1 GI:21176210
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 864)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapps-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM13500 row: e column: 18
High quality sequence stop: 658.

FEATURES
source
location/Qualifiers
1..864

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6156281"
/clone_lib="NIH_MGC_71"
/tissue_type="Telomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 249 a 171 c 231 g 213 t
ORIGIN

Alignment Scores:

Pred. No.:	7.15e-90	Length:	864
Score:	822.00 <td>Matches:</td> <td>170</td>	Matches:	170
Percent Similarity:	82.93%	Conservative:	0
Best Local Similarity:	82.93%	Mismatches:	2
Query Match:	92.78%	Indels:	33
DB:	14	Gaps:	1

US-09-820-003a-2 (1-173) x BQ437146 (1-864)

```
OY 1 MetSerSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuIleGlyAspSer 20
DB 184 ATGTCCAGCATGATGCCGGAATATGATTATTATTCAGTTACTTCTGATTGGCGACTCA 243
OY 21 GlyValGlySerSerCysLeuLeuLeuArgPheAlaAspAspThrTyrThrGluSerTyr 40
DB 244 GGGGTGGGAAGTCTCCCTTCCTTAGGTTGCAGATGATACATATACAGAAACTAC 303
OY 41 IleSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 304 ATACGACCAATTGGTGTGATTCCAATAAAGACATATAGAGTTAGCGGGAACAATC 363
OY 61 LysLeuGlnIle----- 64
DB 364 AACCTTCAATATATGGACACAGCAGGCCAGAAAGATTGCAACAATCACCCTCCAGTTAT 423
OY 65 -----GluserPheAsn 68
DB 424 TACAGAGAGCCCATGGCATCATAGTTGTGTATGATGTACAGATCAGAGTCCCTTCAT 483
```

Mon Nov 18 08:42:06 2002

us-09-820-003a-2.fst

Page 7

OY 69 AasnVallysgIntPrleungIngluileasparytyralasergluasnValasnlyseu 88
DB 484 AATGTTAAACAGTGGCTGCAGGAAATAGATGTTATGACGAGAAATGTCACAAATG 543
OY 89 LeuValIGlyAasnlyscysAspLeuThrThrlYsValValaAspTyThrThrlAllys 108
DB 544 TTGGTAGGACCAATGATGATCTGCACCAAGAAAGTATGATGATGATGATGATGATG 603
OY 109 GluPheAlaAspSerLeuGlylleProPheLeuGluThrSerAlaIysAsnAlaThrAsn 128
DB 604 GAATTTGCTGATCCCTTGGAATCCCTTTTGGAAACAGAGTCTAGCAATGCAAGCAAT 663
OY 129 ValGluInserPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
DB 664 GTAGAACACTCTTCATGACGATGAGTGCATGATTAAGAACAGATGCTCCGAGACA 723
OY 149 ThrAlaGlyAlaGluAlaGluLysSerAsnVallysllegInserThrProVallyslInser 168
DB 724 ACAGCTGGTGGCTGCAGAGATGCAATGGTTAAATTCAGAGCTCCAGTCAAGCACTCA 783
OY 169 Gly-GlyGlyCys 172
DB 784 GGTGGAGGTGTC 796
RESULT 9 901 bp mRNA linear EST 21-AUG-2002
LOCUS B0944368
DEFINITION IMAGE:6441874 5', mRNA sequence.
ACCESSION B0944368
VERSION B0944368.1 GI:22359846
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLES Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM13965 row: m column: 11
High quality sequence stop: 662.
Location/Qualifiers
FEATURES
source
1. 901
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:6441874"
/clone_lib="NCI-CORP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; Salt:
Site-2; NCI: Cloned unidirectionally. Primer: Oligo dT.
Library constructed by: Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

Best Local Similarity: 82.13% Mismatches: 2
Query Match: 92.55% Indels: 34
DB: 14 Gaps: 1
US-09-820-003a-2 (1-173) x B0944368 (1-901)
OY 1 MetSerMetAsnProGlyIleAspTyThrThrlYsValValaAspTyThrThrlAllys 20
DB 188 ATGTCAGCATATATCCGAAATGATATATTATTATTAAGTTACTCTGATGCGGATCT 247
OY 21 GlyValIGlyLysSerCysLeuLeuLeuArgPheAlaAspAspThrTyThrGluSerTy 40
DB 248 GGGGTGGAAAGTCTGCTCTCTCTCTAGCTTCAGATGATGATGATGATGATGATG 307
OY 41 IleSerThrIleGlyValaAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
DB 308 ATGACGACAAATGGTGGTGGATTCATTCAGATACGAACTATAGATTAGGGAACAAATC 367
OY 61 LysLeuGluIle-----GluSerPheAsn 64
DB 368 AAGCTACAGATATGGAGACACAGCAGCGCAAGAAATTTGCAACATCACTTCCAGTTAT 427
OY 65 -----GluSerPheAsn 68
DB 428 TACAGAGAGCCATGCGATCATAGTTGTGTATGATGATGACAGATCAGAGATGCTCAAT 487
OY 69 AasnVallysgIntPrleungIngluileasparytyralasergluasnValasnlyseu 88
DB 488 AACGTTAAACAGTGGCTGCAGGAAATAGATGATGATGATGATGATGATGATGATG 547
OY 89 LeuValIGlyAasnlyscysAspLeuThrThrlYsValValaAspTyThrThrlAllys 108
DB 548 TTGGTAGGACCAATGATGATCTGCACCAAGAAAGTATGATGATGATGATGATGATG 607
OY 109 GluPheAlaAspSerLeuGlylleProPheLeuGluThrSerAlaIysAsnAlaThrAsn 128
DB 608 GAATTTGCTGATCCCTTGGAATCCCTTTTGGAAACAGAGTCTAGCAATGCAAGCAAT 667
OY 129 ValGluInserPheMetThrMetAlaAlaGluIleLysLysArgMetGlyProGlyAla 148
DB 668 GTAGAACACTCTTCATGACGATGAGTGCATGATTAAGAACAGATGCTCCGAGACA 723
OY 149 ThrAlaGlyAlaGluAlaGluLysSerAsnVallysllegInserThrProVallyslInser 168
DB 728 ACAGCTGGTGGCTGCAGAGATGCAATGGTTAAATTCAGAGCTCCAGTCAAGCACTCA 787
OY 168 erGlyGlyGlyCys 173
DB 788 CAGGTGGAGGTGTC 804
RESULT 10 1105 bp mRNA linear EST 05-FEB-2002
LOCUS BM452262
DEFINITION IMAGE:6386191 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:552635
ACCESSION BM452262
VERSION BM452262.1 GI:18501302
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLES Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC/DCID/DTF
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

Mon, Nov 18 08:42:06 2002

us-09-820-003a-2.1st

Page 9

```

OY 109 GIUPEALASPSEITLGLYILEPPOHELEUGLUTHRISERIALALYSASNAATHRASN 128
    |||||||
DB 451 GAAATTCCTATCCCTTGGAATTCCTTTTGGAAACAGCTGCTAGAAATCAACGAAAT 510
    |||||||
OY 129 VALGIUGINSERPEHEMETHMETALALAGLUILELYSLYSARGMETGLYPROGLYALA 148
    |||||||
DB 511 GTAGAACAGCTTTTCATGACGATGCGACGATTAATAAAAGCAATGGAGCA 570
    |||||||
OY 149 THRAGLIGLYALAGLUIYSERASVALYSILEGINSETRHR-PROVALYSGINSE 168
    |||||||
DB 571 ACAGCGGTGTGCTGGAATAATCCAAATGTAAATTCAGACAGCTCCAGTCAGAGCTC 630
    |||||||
OY 168 TGYGLY 170
    |||||||
DB 631 AGGTGGG 637

RESULT 12
AL530265 945 bp mRNA linear EST 13-PEB-2001
LOCUS AL530265 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD009YH19 5
DEFINITION prime, mRNA sequence.
ACCESSION AL530265
VERSION AL530265.1 GI:12793758
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 Evry cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
            Location/Qualifiers
                1..945
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_id="LTI_NFL001_NBC4"
                /sex="male"
                /tissue_type="neuroblastoma cells"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-stranded cDNA was digested with Not I and
                cloned into the Not I and Eco RV sites of the PCMVSPORT 6
                vector. Library was normalized. Pong liang life
                by life technologies. Contact: Pong liang life
                technologies, a division of 20850 USA Fax : (1) 301 610
                8371 Email : filang@lifetech.com URL :
                http://fulllength.lifetech.com"

BASE COUNT 280 a 186 c 224 g 246 t
ORIGIN
Alignment Scores: 3.19e-88 length: 945
Pred. No.: 809.00 matches: 171
Score: 83.01% mismatches: 0
Percent Similarity: 83.01%
Best Local Similarity: 91.31%
Query Match: 91.31% indels: 34
DB: 9 gaps: 1

US-09-820-003a-2 (1-173) x AL530265 (1-945)
OY 1 MetSerSerMetAsnProGluThrAspTyrLeuPheLysLeuLeuLeuLeuLysPse 20
    |||||||
DB 121 ATGTCCACGACGAAATCCGAAATGATGATTTATTTCAAGTACTGATGCGCACATCA 180
    |||||||
OY 21 G1YVALG1LYSERCYSLEULEULEUARGPHEALASPSPTRHTYRTHGLUBERTYR 40
    |||||||

DB 181 GGAGTGGGAAGAGCTTCCCTCTTTAGCTTTCAGAGTACATGATACAGAAAGCTAC 240
    |||||||
OY 41 T1ESERTRH1LEGLYVALASPHELYSL1EARGTHR1LEGL1LEUASPGLYSPTRH1LE 60
    |||||||
DB 241 ATGACGACAAATGGGTGTGATTTCAAAATAGACATRTAGACTTAGCGGGAACAATC 300
    |||||||
OY 61 LYSLEUGL1LE----- 64
DB 301 AACCTTCAAATATGGGACACAGCAGCGCAAGAAAGATTTCGAACATCACTCAGTAT 360
    |||||||
OY 65 ----- GluSerPheasn 68
DB 361 TACAGAGAGCCCATGCAATCAATGTTGTGATGATGTACAGATCAAGATCCTTCATAT 420
    |||||||
OY 69 ASNVALLYSG1NTRPLEUGLNU1LEASPARGTYRALASERGLVASNVALYSLEU 88
    |||||||
DB 421 AATGTTAAACAGTGTCTGAGAAATGATGCTTTATGCCAGTGAATAATGTCACAAATTC 480
    |||||||
OY 89 LEUVALG1LYASNLYSCYASPLEUWTRHTRHLYSYVALYVALASPTYRTHR1ALALYS 108
    |||||||
DB 481 TTGCTAGGGAACAAATGATGATGACCAAGAAAGATGATGACTACACACAGCGAG 540
    |||||||
OY 109 GIUPEALASPSEITLGLYILEPPOHELEUGLUTHRISERIALALYSASNAATHRASN 128
    |||||||
DB 541 GAAATTCCTATCCCTTGGAATTCCTTTTGGAAACCACTGCTAAGAAATCAACGAAT 600
    |||||||
OY 129 VALGIUGINSERPEHEMETHMETALALAGLUILELYSLYSARGMETGLYPROGLYALA 148
    |||||||
DB 601 GTAGAACAGCTTTTCATGACGATGCGACGATTAATAAAAGCAATGGAGTCCGGA-GCA 659
    |||||||
OY 149 THRAGLIGLYALAGLUIYSERASVALYSILEGINSETRHR-PROVALYSGINSE 168
    |||||||
DB 660 ACAGCGGTGTGCTGGAATAATCCAAATGTAAATTCAGACAGCTCCAGTCAGAGCTCA 719
    |||||||
OY 169 -G1YGLYGLYSCYS 173
    |||||||
DB 720 GGGTGGAGGTCTCTCC 735

RESULT 13
LOCUS B073109 642 bp mRNA linear EST 11-DEC-2001
DEFINITION B073109 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone X1108p01 5', mRNA sequence.
ACCESSION B073109
VERSION B073109.1 GI:17503298
KEYWORDS EST.
ORGANISM African clawed frog.
            Xenopus laevis
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Metazoa;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
            Xenopus.
            Xenopus laevis.
            1 (bases 1 to 642)
            K13yAm, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara
            Y.,
            Expressed genes in X. laevis embryo
            Unpublished (2001)
            Contact: Tadasi Shin-I
            National Institute of Genetics
            Center for Genetic Resource Information
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genes.nig.ac.jp.
            Location/Qualifiers
                1..642
                /organism="Xenopus laevis"
                /db_xref="taxon:8355"
                /clone_id="X1108p01"
                /clone_lib="NIBB Mochii normalized Xenopus tailbud
                library"
                /tissue_type="whole embryo"
                /dev_stage="stage 25"

FEATURES
    source

```


Mon Nov 18 08:42:06 2002

us-09-820-003a-2.fst

Page 11

Db 626 GGAATTCGAGATTCCTTGGAAATTCATTTTGGAAACAGCTCTTAGACAGCAGAA 685
Oy 128 nValGluInserPheMetThrMetAlaAlaGluLeuLysArgMet-GlyProGlyA 148
Db 686 TGTAGACAGCTTTTCATGACAGCTGAGATTAAGAGCATGGCTCTGAG 745
Oy 148 lArThrAlaGlyAlaGlyAlaGlyLys-SerAsnValLysIleGlnSerThrProValLysGln 167
Db 746 CTACAGCTGCTGTGGTGGAGAGGCTCAATGTTAAATCAGACACTCCAGTCAAGAG 805
Oy 168 SerGlyGlyGlyCys 172
Db 806 TCAGGTGAGGCTGC 820
RESULT 15
Bi739625 855 bp mRNA linear EST 20-SEP-2001
LOCUS 603361882P1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:536876 5',
DEFINITION mRNA sequence.
ACCESSION Bi739625
VERSION Bi739625.1 GI:15716638
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 855) NIH_MGC_94 /
AUTHORS NIH_MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Stausberg, Ph.D.
Email: cgsab@remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Prepared by: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
CDNA distribution: MGC clone distribution information can be
found through the NIH_MGC Consortium/LNL at:
http://image.llnl.gov
Plate: LLNL1938 row: d column: 21
High quality sequence stop: 849.
FEATURES
source 1..855
location/Qualifiers
organism="Mus musculus"
abxref="taxon:10090"
clone="IMAGE:536876"
clone_1lb="NIH_MGC_94"
tissue_type="retina"
lab_host="DH10B (phage-resistant)"
note="Organ: eye; Vector: PCWV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 234 a 181 c 233 g 205 t 2 others
ORIGIN
Alignment Scores:
Pred No.: 6,36e-88 Length: 855
Score: 806.00 Matches: 169
Percent Similarity: 82.21% Conservative: 2
Best Local Similarity: 81.25% Mismatches: 2
Query Match: 90.97% Indels: 35
DB: 13 Gaps: 1
US-09-820-003a-2 (1-173) x Bi739625 (1-855)
Oy 1 MetSerMetAsnProGluTyrAspTyrLeuPheLysLeuLeuLeuIleGlyAspSer 20
Db 185 ATGTCCAGCATGATCCGAAATGATGATTTATTCAGATCTCTGATGGGATTC 244
Oy 21 GlyValGlySerCysLeuLeuLeuArgPheAlaAspThrTyrThrGluSerTyr 40

Db 245 GGGGTGGAAAGTCTGCTCTCTCTTGTAGGTTCAGATGATAGCTATGCGAAAGCTAC 304
Oy 41 lIeSerThrIleGlyValAspPheLysIleArgThrIleGluLeuAspGlyLysThrIle 60
Db 305 ATCAGACAAATGGTGTGGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 364
Oy 61 LysLeuGlnIle----- 64
Db 365 AACCTACAGATTTGGAGCAGACAGAGGCCAGAGAAAGATTTCGAACATCCTCCAGTAT 424
Oy 65 ----- GluSerPheAsn 68
Db 425 TACAGAGAGCCATGCCATCATATGTTGTATGATGTACAGATTCAGAGATCTTCAT 484
Oy 69 AsnValLysGlnTTPLeuGlnGluIleAspArgTyrAlaSerGluAsnValAsnLysLeu 88
Db 485 AACGTTAACAGAGTGTGCTGACAGAGATAGATGCTACGCCAGTGAATAATGTCACAGAT 544
Oy 89 LeuValGlyAsnLysCysAspLeuThrThrLysLysValValAspTyrThrThrAlaLys 108
Db 545 GTGGTAGGAGAACAAATGTGACTGACCAAGAAAGTAGAGTACACAAACAGCAAG 604
Oy 109 GluPheAlaAspSerLeuGlyIleProPheLeuGluThrSerAlaLysAsnAlaThrAsn 128
Db 605 GAATTTGCAATTCCTTGCATTCATGATTCATTTTGGAAACAGTGTCAAGAACAGAT 664
Oy 129 ValGluInserPheMetThrMetAlaAlaGluLeuLysLysArgMetGly--ProGlyA 148
Db 665 GTAGAACGCTCTTCATGACAGTGCAGATTAAGGAAATGGTCCCTGAG 724
Oy 148 lArThrAlaGlyAlaGlyAlaGlyLysSerAsnValLysIleGlnSerThrProValLysGln 168
Db 725 CTACAGCTGCTGTGGTGGAGAGTCCAAATTTAAATCAGAGCATCCAGTCAAGCAGT 784
Oy 168 SerGly-GlyGlyCysCys 173
Db 785 CAGGTGAGGCTGCTGC 802
Search completed: November 17, 2002, 14:30:41
Job time : 1564 secs

